



Review

# Enhancing Clinical Utility: Utilization of International Standards and Guidelines for Metagenomic Sequencing in Infectious Disease Diagnosis

Chau-Ming Kan <sup>1</sup>, Hin Fung Tsang <sup>1</sup>, Xiao Meng Pei <sup>2</sup>, Simon Siu Man Ng <sup>3</sup>, Aldrin Kay-Yuen Yim <sup>4</sup>, Allen Chi-Shing Yu <sup>4</sup>  and Sze Chuen Cesar Wong <sup>2,\*</sup>

<sup>1</sup> Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong, China; kantrevor@gmail.com (C.-M.K.); andy\_thf@yahoo.com.hk (H.F.T.)

<sup>2</sup> Department of Applied Biology & Chemical Technology, The Hong Kong Polytechnic University, Hong Kong, China; xiaomeng.pei@connect.polyu.hk

<sup>3</sup> Department of Surgery, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China; simon.ng@cuhk.edu.hk

<sup>4</sup> Codex Genetics Limited, Shatin, Hong Kong, China; aldrinyim@codexgenetics.com (A.K.-Y.Y.); allenyu@codexgenetics.com (A.C.-S.Y.)

\* Correspondence: cesar.wong@polyu.edu.hk

**Abstract:** Metagenomic sequencing has emerged as a transformative tool in infectious disease diagnosis, offering a comprehensive and unbiased approach to pathogen detection. Leveraging international standards and guidelines is essential for ensuring the quality and reliability of metagenomic sequencing in clinical practice. This review explores the implications of international standards and guidelines for the application of metagenomic sequencing in infectious disease diagnosis. By adhering to established standards, such as those outlined by regulatory bodies and expert consensus, healthcare providers can enhance the accuracy and clinical utility of metagenomic sequencing. The integration of international standards and guidelines into metagenomic sequencing workflows can streamline diagnostic processes, improve pathogen identification, and optimize patient care. Strategies in implementing these standards for infectious disease diagnosis using metagenomic sequencing are discussed, highlighting the importance of standardized approaches in advancing precision infectious disease diagnosis initiatives.

**Keywords:** infectious disease diagnosis; next generation sequencing; metagenomics sequencing; implementation and validation; ISO15189; ISO24420; ISO20397



**Citation:** Kan, C.-M.; Tsang, H.F.; Pei, X.M.; Ng, S.S.M.; Yim, A.K.-Y.; Yu, A.C.-S.; Wong, S.C.C. Enhancing Clinical Utility: Utilization of International Standards and Guidelines for Metagenomic Sequencing in Infectious Disease Diagnosis. *Int. J. Mol. Sci.* **2024**, *25*, 3333. <https://doi.org/10.3390/ijms25063333>

Academic Editor: Stephen Bustin

Received: 2 February 2024

Revised: 11 March 2024

Accepted: 12 March 2024

Published: 15 March 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Infectious diseases caused by viruses, bacteria, parasites, or fungi, such as hepatitis B and C, diarrheal diseases, HIV/AIDS, tuberculosis (TB), respiratory infections, pneumonia, condylomas, etc., can be transmitted directly or indirectly from one person to another [1]. For many years, medical research and public health efforts have focused on identifying and managing infectious diseases. However, the understanding and management of infectious diseases have evolved. In the 20th century, molecular diagnostic technologies, such as polymerase chain reaction (PCR) and other DNA-based methods, such as in situ hybridization and Southern blot analysis, were used to identify specific RNA or DNA sequences of infectious agents and revolutionized the field of infectious disease diagnosis [2,3]. The utilization of next-generation sequencing (NGS) technology, also known as metagenomics sequencing (mNGS), in combination with bioinformatics, has also enabled researchers to identify pathogenic microorganisms more rapidly and accurately, compared with the conventional approaches, such as culture-based methods and serology methods like enzyme-linked immunosorbent assay (ELISA) [4–6]. In addition, genomic analyses facilitated by NGS can offer valuable insights for viral quasispecies analysis [7–12],

and for genomics-informed outbreak investigations, including epidemic dynamics, source, and timing, and the geographical origin of pathogens [13–20]. By analyzing the genetic information obtained through NGS, researchers can gain a deeper understanding of the transmission patterns and genetic diversity of infectious agents, aiding in the development of effective control and prevention strategies [21–24].

Clinical mNGS has proven to be an effective diagnostic tool for infectious diseases due to its ability to detect a large range of pathogens. Recent advances in technology have made it an invaluable diagnostic tool for clinicians. mNGS is particularly useful for detecting infectious agents in cases of encephalitis and meningitis [25,26]. As an example, some studies have demonstrated the potential of mNGS in the diagnosis of meningitis and encephalitis. Fan et al. have demonstrated that by utilizing mNGS, infectious diseases caused by pathogens that are less frequently encountered, such as *Angiostrongylus cantonensis*, can be accurately diagnosed, which in turn results in improved patient outcomes and care [26]. Wilson et al. also demonstrated the comprehensive spectrum of potential causes that can be identified by a single mNGS assay, including viral, bacterial, fungal, and parasitic pathogens, showcasing its utility in diagnosing complex central nervous system infections [25]. Clinical mNGS has also shown promise in diagnosing various infectious diseases, including tuberculosis, HIV, and coinfections [27]. Wang et al. (2019) explored the potential of mNGS in identifying pathogens causing tuberculous meningitis, highlighting its potential to enhance tuberculosis-related infection diagnosis [27]. In addition, Liang et al. (2023) demonstrated the utility of mNGS in identifying multiple pathogens simultaneously by detecting coinfections with SARS-CoV-2 and influenza A [28]. As well as this, Xie et al. (2023) demonstrated that mNGS provided a sensitive and unbiased method for detecting multi-pathogenic pneumonia in HIV-infected patients [29]. These studies demonstrate mNGS's versatility and effectiveness in diagnosing a variety of infectious diseases, providing clinicians with a comprehensive and advanced diagnostic tool.

While mNGS holds immense potential for a wide range of applications, implementing it into clinical settings for diagnosing infectious diseases poses several challenges. The challenges associated with the standardization of workflows in mNGS have a significant impact on the turnaround time and costs [30,31], data quality, reproducibility [32], comparability [33], and biological interpretation [34–36]. Therefore, recent studies have made significant progress in standardizing metagenomics workflows both in the wet laboratory (wet lab) [37,38] and the dry laboratory (dry lab) [39–41]. For instance, Parker et al. (2023) developed a sample-to-answer workflow called PanGIA that includes simplified, standardized wet lab procedures and data analysis with an easy-to-use bioinformatics tool [38]. In addition, the International Organization for Standardization (ISO) has also published international standards for the data processing of shotgun metagenomic sequences (ISO24420:2023) [42] and for the workflow and quality evaluation of sequencing data from massively parallel sequencing (ISO20397-1:2021 and ISO20397-2:2021) [43,44]. In this review, different mNGS approaches and challenges in infectious disease diagnosis will be firstly discussed, along with the strategies and validation of mNGS, and then implementing it into clinical settings will be discussed based on different published guidelines and international standards.

## 2. Understanding the Role of ISO Standards in Workflow Standardization and Different mNGS Approaches

The operation of mNGS is mainly divided into two areas, which are the bioinformatics pipeline level (dry lab) and then the entire mNGS pipeline level (wet lab). A well-established protocol for experiment design, starting from sample collection, processing, nucleic acid extraction, library preparation, and sequencing in the wet lab, to binning, assembly, annotation, and visualization in the dry lab, is imperative to ensure reproducibility and accuracy of mNGS workflows. An integrated quality management approach, in line with regulatory standards, is required for any NGS-based implementation in the clinical laboratory [45]. The wet lab workflow and quality evaluation of sequencing data in the

dry lab section should meet ISO20397 standards. The data processing specific to mNGS should follow the general requirement under the ISO24420 standards. To be used in clinical laboratories, assays should follow ISO15189 [46,47], a global quality standard for use in clinical laboratories, and standardized standard operating procedures (SOP). Changing any component of an assay should be validated and meet acceptable performance criteria before testing on patients, including reagents, equipment, sequencing instruments, and bioinformatics tools. Continual updating of the SOP, the reagent manual, the equipment calibration, and participation in external quality assessment (EQA) schemes are necessary to incorporate intermediate technological advancements. Besides that, laboratories must define clearly the intended clinical use for unbiased metagenomic testing, as well as the types of pathogens that will be detected and reported. There are many factors to be considered, including turnaround times, specimen requirements, library preparation protocols, sequencing platforms, sequencing depth, quality control, data analysis, costs, and clinical relevance [48]. Results review and reporting workflow should be clearly defined by laboratories, so they can choose the right data analysis solutions [48]. It is essential to be familiar with all available options, along with upcoming protocols for validation, given that many workflow components starting from the sample collection to reporting are constantly being developed and optimized.

An important challenge of metagenomic pathogen detection is the validation of an assay which cannot be declared validated until it has been applied to a wide variety of microorganisms almost endlessly. Assays that are not validated may generate incorrect results, leading to incorrect clinical interpretations and treatment decisions. There are no definitive recommendations for the clinical implementation of mNGS testing, even though the US Food and Drug Administration (FDA) has provided guidelines for clinical validation of NGS-based infectious diseases testing [49]. It is therefore recommended to learn from failed experiments and to use representative organisms to evaluate performance characteristics, followed by ongoing monitoring of assay performance and independent confirmation if unexpected results occur.

### *2.1. Understanding the Current mNGS Strategies and Challenges for Infectious Disease*

A number of mNGS approaches have shown promise for clinical diagnosis of infectious diseases based on metagenomics, including short-read sequencing, long-read sequencing, and RNA sequencing.

#### *2.1.1. RNA Sequencing*

In metagenomics, RNA sequencing is useful for revealing the active gene expression profiles of microorganisms, thereby revealing the functional activities of the microbes and their responses to their environments [50]. Gene expression profiles within complex microbial communities can be comprehensively analyzed using this approach, providing insights into the functional activities and responses of microorganisms [51]. A major advantage of RNA sequencing in metagenomics is its ability to capture real-time functional dynamics of microbial communities, thus allowing the monitoring of gene expression changes in various environments [51]. Additionally, RNA sequencing can also help detect novel RNA viruses and provide insight into RNA viral diversity and evolution in metagenomic samples [52]. However, RNA sequencing in metagenomics poses challenges, such as the inherent instability of RNA molecules, which results in transcript degradation and biases in transcript representation [53]. This instability can lead to inaccurate gene expression quantifications and data analysis artifacts [53,54]. Moreover, microbial RNA samples containing host-derived RNA can complicate the analysis and interpretation of metagenomic data; therefore, robust bioinformatics pipelines are required for precise taxonomic and functional analysis [54].

### 2.1.2. Long-Read Sequencing

Long-read sequencing in metagenomics involves the usage of sequencing technologies that produce longer DNA reads, such as those provided by Pacific Biosciences and Oxford Nanopore, for analyzing the genetic contents of microbial communities present in environmental samples. This approach offers several advantages in metagenomic studies. In comparison to short-read technologies, long-read sequencing enables a more accurate reconstruction of the bacterial genome by enabling the assembly of complete bacterial genomes with a greater degree of contiguity [55]. Longer reads also enable better bridging of repetitive sequences within and between genomes, resulting in more complete genome assemblies [56]. Further, long-read sequencing improves the accuracy of metagenomic annotations and profile recovery by allowing complete genes, operons, and repetitive elements to be retrieved before assembly [57]. Despite its advantages, long-read sequencing in metagenomics also presents challenges. The higher cost of generating long reads compared to short reads may limit the adoption of long-read technologies for metagenomic studies [58]. Furthermore, long-read sequencing may suffer from high DNA requirements and sequencing errors, which can affect accurate gene prediction and assembly [56]. Moreover, long-read sequencing for metagenome assembly may be limited by limitations such as incomplete databases, insufficient read lengths, and high sequencing error rates [58].

### 2.1.3. Short-Read Sequencing

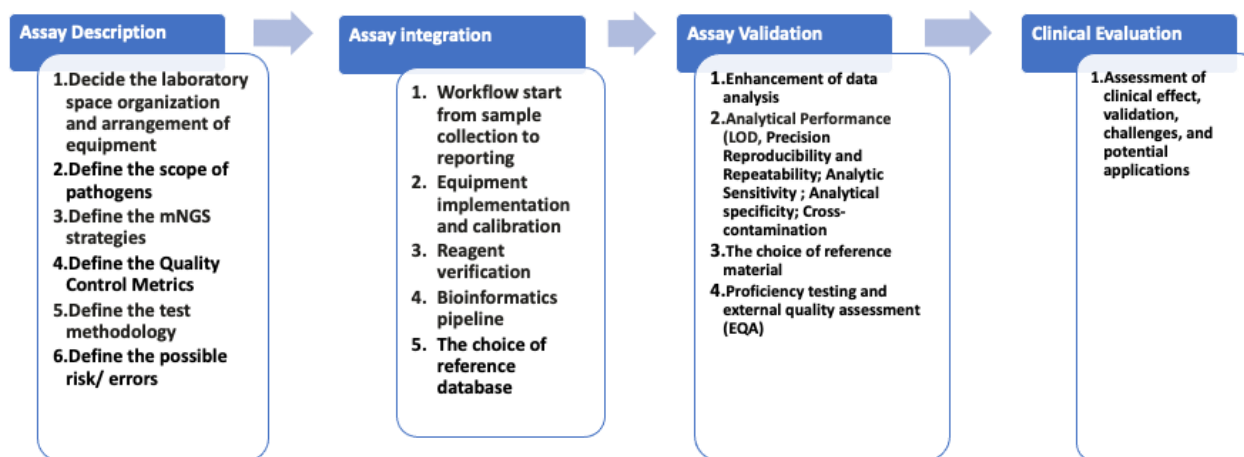
Short-read sequencing in metagenomics typically generates DNA sequences that are relatively short, often around 50 to 300 base pairs. There are several advantages and challenges associated with this method, which is typically used by platforms like Illumina and SOLiD. Some major advantages are greater throughput, lower cost, and well-established data analysis pipelines [59], and the ability to target specific regions of interest using enrichment methods such as amplicon sequencing [60,61] or capture-based techniques [62,63]. As a result of targeted sequencing, low-abundance pathogens within complex microbial communities can be detected more sensitively, improving diagnostic capabilities [64]. In addition, short-read sequencing with enrichment can provide higher coverage depth, allowing for greater accuracy when it comes to the identification of microbial species and genes for antimicrobial resistance [65]. As an example, Illumina amplicon sequencing enables characterization of bacterial communities at a high resolution by targeting specific regions of the 16S ribosomal RNA gene, providing valuable information about the diversity and composition of bacterial communities [60,61]. Using microbiota from the gut, soil, and bacterial infections has been crucial to the study of various ecosystems, providing valuable insights into microbial populations and their interactions [60].

However, short-read sequencing can also pose challenges in metagenomics. One significant limitation is the difficulty in assembling complex metagenomes due to the short lengths of the reads, which can lead to fragmented assemblies [66]. It is especially difficult to reconstruct individual genomes in mixed microbial communities when multiple genomes are present [67]. Additionally, short-read sequencing may encounter repetitive regions and structural variations, making genome assembly less accurate and complete [68]. For the target sequencing, one key challenge is the risk of amplification bias during enrichment, which causes the data to reflect a skew of certain microbial taxa or genomic regions, which can affect microbial diversity quantification and prevent detection of rare or novel organisms [69].

## 3. Idea of Metagenomics Sequencing Assay Implementation Workflow

Implementing metagenomics sequencing assays for infectious disease diagnosis involves several key considerations. Currently, the FDA is actively assessing the components necessary for the approval of NGS-based tests for infectious diseases, which are primarily used to detect antimicrobial resistance and identify pathogens [70]. The FDA released guidelines for designing, developing, and validating next-generation sequencing tests, paving the way for routine NGS-based testing in clinical laboratories [71]. In the following

sections, we will discuss how the workflow of mNGS implementation can be divided into several steps, based on various studies, guidelines, and international standards, including assay description and definition, workflow integration and optimization, assay validation, and clinical evaluation (Figure 1).



**Figure 1.** Overview of metagenomics sequencing assay implementation. Before implementing in the clinical setting, the scope of the pathogen must be defined to decide the choice of sample type, the need for equipment and reagents, and the test methodology. An error-based approach should be used before assay validation to find the potential risks in the procedure. Clinical evaluation with real samples is important before implementing into the clinical setting.

#### 4. Assay Description

Assay description and definition are crucial components in the process of assay integration, as they provide a clear understanding of the assay methodology and parameters before its integration into a real system and the establishing of standardized protocols. According to the guidelines from the FDA, “Infectious Disease Next Generation Sequencing Based Diagnostic Devices: Microbial Identification and Detection of Antimicrobial Resistance and Virulence Markers”, because of the large diversity of specimen types and infectious disease agents that can be present in samples, straightforward pre-analytical, biochemical, or bioinformatics processes are not allowed. Each unique specimen type may require a different protocol for nucleic acid extraction and library preparation, as well as a unique bioinformatics algorithm to generate the final clinical result; therefore, the completed assay description needs to be defined clearly [49].

##### 4.1. Deciding on Laboratory Space Organization and the Arrangement of Equipment

Since mNGS is unbiased and there is an increased risk of detection and interference caused by cross-contamination, proper laboratory organization and diagnostic workflow are essential [37]. Workflow must be unidirectional between the pre-amplification and post-amplification areas. Reagent preparation rooms must be clean, free of samples, and separated from post-amplification and pre-amplification areas to avoid any contamination of the reagent and the mastermix. A laboratory’s equipment, such as freezers, refrigerators, pipettes, safety cabinets, thermocyclers, vortexers, centrifuges, etc., should be clearly labeled and assigned to different rooms. All equipment must be calibrated by accredited bodies and the calibration intervals of all equipment should be based on the ISO15189 standards.

##### 4.2. Defining the Scope of Pathogens

The laboratory should first define the pathogens that will be detected by the assay, based on the intended use, the type of specimen, and the characteristics of the patient. The choice of pathogens influences the choice of extraction method, sequencing strategies,

target enrichment or host nucleic acid depletion, reference database, sequencing depth and coverage, and data analysis tools. Besides that, there are several factors to consider when determining the scope of pathogens. These factors include genotype characteristics, environmental conditions, and non-random factors that influence pathogen damage to varieties [72]. Additionally, the transmission dynamics of pathogens can be influenced by challenges encountered in zoonotic disease surveillance, insufficient spatial information, and a lack of understanding of surveillance scope [73]. Emerging infectious diseases are significantly correlated with socioeconomic, environmental, and ecological factors, providing a way to identify regions that are likely to develop new diseases [74]. An integrated approach to zoonotic spillover requires an alignment of ecological, epidemiological, and behavioral factors influencing pathogen exposure, as well as factors influencing susceptibility to infection within humans [75]. Furthermore, the vector-to-host ratio plays a pivotal role in determining the severity of pathogen transmission and infection risk [76]. Moreover, microbial pathogens' vulnerability affects their resistance and pathogenicity, making them crucial in determining the scope of pathogens [77]. Factors such as human connectivity, increasing antimicrobial resistance, and dynamic human behavior influence prevention and control, affecting the scope of pathogen dynamics [78].

#### 4.3. Defining the mNGS Strategies

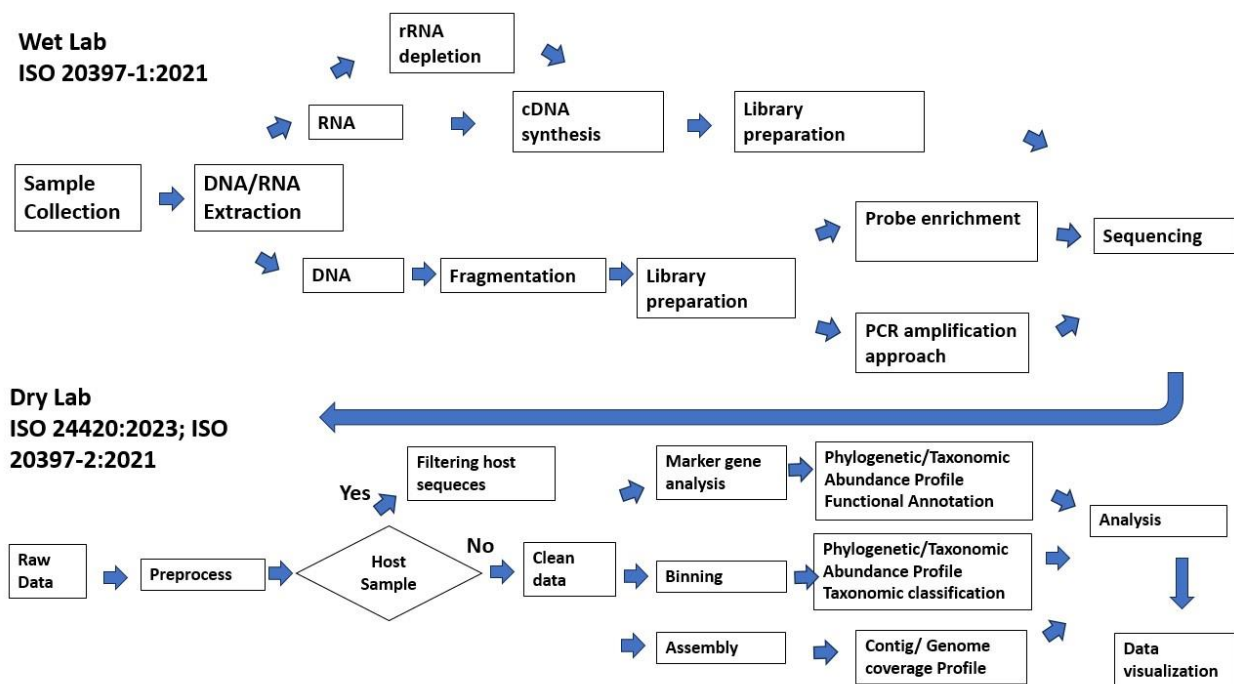
mNGS strategies should be determined according to the application before implementing mNGS assays. To improve the detection and analysis of microbial DNA in complex samples, mNGS strategies primarily use sensitivity and enrichment or depletion methods. mNGS has a major limitation with the presence of a high background, either from the host or from the microbiome. As an example, Wilson et al. evaluated the clinical effect of mNGS in diagnosing meningitis and encephalitis, highlighting the limitations associated with low sensitivity when the human host or microbiome background is high [25]. Furthermore, Salipante et al. were able to show the limitations of low sensitivity due to high background from the microbiome when they compared the performance of Illumina and Ion Torrent sequencing platforms for bacterial community profiling [79]. Additionally, the importance of monitoring low-biomass samples for background contamination was also highlighted by Weyrich and colleagues [80].

To efficiently perform mNGS and improve microbial DNA detection in different samples, a host depletion approach is crucial. A variety of approaches have been used to overcome host contamination, including capture probes for subtractive hybridization [21,81], CRISPR-Cas9 cleavage on target sequence [82] and ribonuclease (RNase) H-based depletion methods [83], as well as nanopore adaptive sequencing [84]. In the case of RNA libraries, DNase I treatment should be performed after extraction to remove residual human background DNA [85]. Aja-Macaya et al. have recently developed an efficient method for enriching the mNGS of monkeypox viruses using saponin/NaCl combination DNase treatments, combined with high g-force centrifugation, resulting in 96% of reads being classified as not human DNA using the enrichment method, and 5–10% using the non-enrichment method [86].

#### 4.4. Defining the Test Methodology

A complex and computationally intensive process is involved in transforming raw sequence data into actionable information in NGS, starting with sample collection, extraction, fragmentation, library preparation, sequencing, data analysis, and interpretation of the results (Figure 2). As a brief overview, the DNA/RNA are first extracted from the samples. Following the extraction of DNA/RNA, a library is constructed where the DNA of each organism is sheared into fragments and inserted into adapters containing barcodes to allow multiplexing of hundreds of samples based on the quality criteria for the selected NGS platforms and applications. Using a preferred NGS technology, these individual libraries are pooled together and submitted for analysis. After the sequencing has been completed, several bioinformatics steps need to be performed, including quality control,

read alignment or assembly, variant calling, and annotation, which is dependent on the specific application [87]. An analysis of the assembled genome against reference strains facilitates the identification of pathogens, strain typing at high resolution, and prediction of important phenotypic characteristics (e.g., virulence, resistance to antibiotics) [88]. There is a need for well-curated and up-to-date reference databases since microorganisms are constantly evolving and plasmids encoding traits of virulence and antimicrobial resistance can be exchanged across strains and species. By comparing assembled genomes with others, we can look for evidence of transmission based on phylogenetic clustering [88]. Different bioinformatics tools are required for each step, including assembly, strain typing, phenotyping, and clustering. Therefore, bioinformatics expertise and computational resources are required to perform the analysis using various software tools and pipelines [87,88].



**Figure 2.** Overview of different sequencing approaches for infectious disease diagnosis. The wet lab workflow is based on ISO20397-1:2021, which included sample collection, DNA/RNA extraction, library preparation, target capture, sequencing, and bioinformatics analysis. The dry lab workflow is based on ISO24420:2023 and ISO20397-2. Three main metagenomic analysis approaches, including marker gene analysis, binning, and assembly, are described in this figure.

#### 4.5. Defining the Possible Risk/Errors in the Whole Workflow 5M1E Methodology

The 5M1E methodology (manpower, method, machine, material, measurement, and environment) was mainly involved in the process of the root cause analysis to enhance problem-solving capabilities, optimize procedures, and maintain high-quality assurance standards. A systematic approach, the Plan–Do–Check–Act (PDCA) cycle, was also incorporated into the 5M1E methodology, which involves four stages: planning, implementing, evaluating, and action-taking based on the evaluation. Therefore, the 5M1E methodology needs to be applied to the whole process, starting from pre-implementation to implementation steps, including assay design, quality control matrix determination, and assay validation. An overview of the subject, action, and objective based on the 5M1E methodology using an NGS assay is shown in Table 1. The 5M1E methodology can also be applied in molecular laboratories. The potential reasons for errors are listed in Table 2.

**Table 1.** Overview of the subject, action, and objective based on the 5M1E methodology using on NGS assay.

	Manpower	Machine	Material	Method	Environment	Measurement
<b>Subject</b>	Well-trained personnel, compliance with SOP	Automation extraction machine	Raw materials, reagent and samples	Procedures, method, and protocols	Environment conditions such as humidity, lighting, temperature, and cleanliness	Instrument, techniques, data, and tools
<b>Action</b>	Training, skill development, and monitoring of ongoing performance	Calibration, routine maintenance; adhering to equipment specification	Verification; sample acceptance criteria, reagent specification checking, storage and handling	Validation	Adequate controls, monitoring	Calibration, validation, and routine performance check
<b>Objective</b>	Minimize human-related variation, human error	Minimize variations in test results; ensure that the equipment is performing within acceptable limits and meets the required specifications; reduces the risk of errors and inaccuracies in test results	Ensures the integrity and traceability of materials, prevents contamination and degradation, and ensures high-quality materials are acquired	Ensure reliable and accurate result	Minimize variations caused by environmental factors	Ensure precise measurement

**Table 2.** An error-based approach using 5M1E methodology on NGS assay.

mNGS Procedure	Error/Risk	Consequence	5M1E	Means of Minimizing Risk
Sample collection	Improper storage [89] and handling of samples, such as higher freezing and thawing cycle	<ul style="list-style-type: none"> <li>Degradation of nucleic acids</li> <li>Inaccurate representation of the microbial community present in the sample [89,90]</li> </ul>	Method, Materials, Environment	<ul style="list-style-type: none"> <li>Use sterile collection tools and containers</li> <li>Proper training of personnel involved in sample collection</li> </ul>
Sample collection; nucleic acid extraction	Host DNA contamination/cross-contamination of samples	<ul style="list-style-type: none"> <li>Influence the accuracy and reliability of sequencing results [91]</li> <li>Reduced sensitivity to detect microbial reads [92]</li> <li>Inaccurate classification of pathogenic communities [93]</li> <li>Misidentification of DNA samples and compromised experimental results [94,95]</li> </ul>	Method, Machine, Materials, Environment	<ul style="list-style-type: none"> <li>Reduce contamination by evaluating host depletion and extraction methods [83]</li> <li>Ensure sequence distribution and individual sequence errors are addressed with quality control measures [96]</li> <li>Calibration and decontamination of the machine</li> </ul>
Nucleic acid extraction	Inadequate lysis of cells or tissues/non-verified reagent	<ul style="list-style-type: none"> <li>Suboptimal DNA yield and compromised quality [97,98]</li> </ul>	Materials	<ul style="list-style-type: none"> <li>Implementing methods that selectively lyse host cells and/or separate specific organisms from clinical samples [99,100]</li> </ul>



Table 2. Cont.

mNGS Procedure	Error/Risk	Consequence	5M1E	Means of Minimizing Risk
Nucleic acid extraction; library preparation	Improper sample preparation, such as incorrect sample loading or inadequate mixing	<ul style="list-style-type: none"> <li>Inaccurate biological interpretations and impact on the reliability of the sequencing data [89,101]</li> <li>Affecting downstream application [102]</li> </ul>	Methods	<ul style="list-style-type: none"> <li>Proper training of personnel involved in the procedure</li> </ul>
Nucleic acid extraction	Inhibitory Substances	Interfere with the downstream analysis and application [103]	Methods	<ul style="list-style-type: none"> <li>Optimization of DNA extraction and purification procedures minimizes co-extraction of inhibitory substances with DNA [104]</li> <li>Evaluate the DNA preservation methods to avoid co-extraction of inhibitory substances [105]</li> </ul>
Library preparation	High levels of DNA fragmentation	<ul style="list-style-type: none"> <li>Reduced DNA yield and inaccurate quantitation, impacting the reliability of subsequent analyses [102,106]</li> </ul>	Methods	<ul style="list-style-type: none"> <li>Evaluate the incubation time for fragmentation</li> </ul>
Library preparation	Index hopping	<ul style="list-style-type: none"> <li>Erroneous conclusions and misinterpretations of the microbial community composition [107]</li> </ul>	Methods	<ul style="list-style-type: none"> <li>Incorporating unique dual-indexing strategies that assign unique index sequences to each sample [108]</li> </ul>
Sequencing	Systematics error/sequencing error	<p>Overestimation of gene and taxon abundance [109]</p> <p>Inaccurate gene prediction on short reads [110]</p>	Machine, Methods	<ul style="list-style-type: none"> <li>A simulation tool that incorporates error models based on explicit errors and coverage bias can help address sequencing errors [111]</li> <li>The development of a statistical method for functional profiling can allow sequencing errors to be estimated and possible functions to be detected in metagenomic data [112]</li> </ul>
Bioinformatics analysis	<p>Misinterpretation of results [113,114]; inadequate data preprocessing [115]; incorrect taxonomic or functional annotation [116]; failure to account for batch effects [117]; inadequate validation of findings [118]; lack of reproducibility [119]; overfitting or underfitting of models [120]; inadequate training data [121]</p>	Impact the accuracy and reliability of the results	Manpower, Machine, Methods	<ul style="list-style-type: none"> <li>Integrate comprehensive databases and update them regularly to incorporate newly discovered microbial species.</li> <li>Developing novel taxonomic classification algorithms and tools [122]</li> <li>Conduct of rigorous quality control measures as well as benchmarking and standardizing bioinformatics workflows</li> <li>Using validated statistical methods, coupled with careful consideration of data normalization and bias correction [123]</li> <li>Proper training of personnel</li> </ul>

#### 4.6. Defining the Quality Control Metrics

Quality control is a critical aspect of mNGS to ensure the accuracy and reliability of the data. It is necessary to establish quality control measures and to apply them to every run of the assay. Specimen acceptability, nucleic acid quantification, and library qualification and quantification should be monitored since libraries with inadequate fragment sizes and underloading can result in reduced coverage [124]. A quality control (QC) process should be carried out on raw mNGS data to ensure that the data are clean. The QC metrics and follow-up actions are summarized in Table 3 following ISO24420:2003, ISO20397-1:2022, Clause 4 and 8.3, ISO20397-2:2021, 4.3, and different metagenomics studies.

**Table 3.** Indicators of validation and QC metrics.

Measurement	Indicators of Validation	QC Metrics	Follow-Up Actions
Nucleic acid quality	DNA Purity	High DNA purity [125]: A260/A280: 1.88–1.94 Low DNA purity: A260/A280 < 1.6	Perform additional purification step, such as phenol-chloroform extraction or silica-column-based purification kits Optimize the extraction protocols to minimize contamination and improve the purity and quantity Perform additional purification step, such as DNase treatment or RNA cleanup kits Re-extraction
	RNA Purity	RNA quality acceptable range [126]: A260/A280: around 2 A260/A230: 2–2.2 RNA integrity number (RIN) analysis depending on different sample types Acceptable range 7–8	
Nucleic acid quantity	DNA/RNA concentration	Depend on the library preparation approaches: DNA: 20ug [127] 50–250 ng [128] RNA [129]: standard (1 ug); low (10 ng–100 ng); ultra-low (<1 ng)	
Library quality and quantity	Library size distribution and quality patterns Library concentration	For short-read sequencing Size distribution: 250–350 For long-read sequencing Size distribution: 10–25 kb (PacBio@HiFi sequencing) [130] 250 bp to 50 kbp (Nanopore sequencers (MinION, GridION, and PromethION) [131]	Ensure that fragment sizes within the narrow range of expected molecular weight match the measured fragment sizes Prevent adapter dimers, primer dimers (~150 bp), and molecular weight outside the expected range
Sequencing quality	Total run yield (Gb), sequencing cluster density (K/mm <sup>2</sup> ), % reads passing filter, % bases ≥Q30, sequencing error rate, sequencing read length	For short-read sequencing: the sample base mass value shall be more than 20 if Q20 is greater than 90%; and the sample base mass value shall be more than 30 if Q30 is greater than 80% [42]	Resequencing should be conducted if sequencing QC cannot be achieved
Coverage depth	Sequence depth should be evaluated before sequence assembly, taking the sample complexity into consideration [42]	A sufficient level of sensitivity and specificity must be achieved in the regions of interest Short read: at least 30× coverage	Reanalysis of samples should be conducted if coverage thresholds exceed the validated range; an alternate method may be used if only local regions are affected

Table 3. Cont.

Measurement	Indicators of Validation	QC Metrics	Follow-Up Actions
GC bias	GC content	No specific threshold for GC bias GC content <70% [132]	Optimization of library preparation [133] Implement filtering strategies [133] GC content correction [134] Evaluation of sequencing bias with different library preparation kit [135] Normalization techniques [136]
Base call accuracy	Phred quality score (Q), where $Q = -10 \log_{10}(P)$ [137]	Good base call quality scores Phred score > 30 [138] Base calling: Short-read sequencing: >99.9% Long-read sequencing: ~90%	Each run should be monitored for quality scores and signal-to-noise ratios The results of low-quality scoring can lead to more false positive variant calls; therefore, repeat testing may be necessary
Duplication rate	Duplication rate	Maximum duplication rate should be defined for each assay.	Optimization of library preparation Adjust PCR conditions [42,43] Unique molecular identifiers (UMIs) [139] Filtering duplicate reads [140] Error correction algorithms [134]
Mapping quality	Mapping quality scores	Map quality parameters must be established during validation in order to filter out reads that map to nontargeted regions (insertion or indel) and uncertain bases (N characters) [42]	The mapping quality of each run must be monitored as non-specific amplification, off-target DNA capture, or contamination may result in poor results

## 5. Assay Integration and Optimization

After defining the assay and describing its components, the next crucial step in the workflow is the integration and optimization of the entire process. This involves streamlining the various stages of the assay, from sample collection to data analysis, to ensure efficiency and accuracy.

### 5.1. Choice of Sample Type

The presence of low pathogen abundance is one of the challenges in detecting and identifying pathogens in complex microbiomes [141]. Microorganisms that live in the environment or serve as commensals may make it more difficult to identify true pathogens [141]. Therefore, it is important to select the right sample type based on the type of microorganisms to be studied. Multiple factors should be considered when determining the appropriate sample type for metagenomic sequencing. There are several factors to consider, including the research objectives, the target microbes, and the resources and technologies available. For instance, if viral genomes are of interest, blood or respiratory samples may be appropriate [142]. In contrast, for bacterial infections, samples from the lower respiratory tract or clinical specimens may be more suitable [143]. Data from mNGS should also be used for the intended purpose. In outbreak investigations, the source of infection and its spread can be determined by sample collection from affected individuals or environmental sources [15,144]. Similarly, contaminated food samples or samples from affected individuals can be tested for pathogens in foodborne outbreak investigations [144]. The use of mNGS for AMR surveillance can be carried out with samples from sewage or wastewater [145], fecal samples [146,147], or respiratory samples, such as nasopharyngeal swabs or sputum samples from human populations. It is possible to obtain valuable information from these types of samples regarding how antimicrobial resistance genes are prevalent, diverse, and transmitted within populations. For instance, fecal samples can provide insight into the gut microbiota and its role in the development and transmission of AMR [147]. The gut microbiota has been altered in patients with COVID-19, and the abundance of specific species of bacteria has been correlated with the severity of the disease [147].

### 5.2. Sample Preparation

Preparation of samples is the first and most important step in any metagenomics analysis. Ideally, metagenomic sampling aims to collect enough biomass to perform sequencing while minimizing contamination. Samples from the same environment can have varying microbial content, making it more difficult to detect statistically significant and biologically relevant differences between them. To ensure sufficient statistical power, it is crucial to design experiments appropriately, especially for small samples [148]. The conditions under which samples are collected and preserved can affect the quality and accuracy of metagenomic data, and they are also important confounders when analyzing data from different studies [148]. In addition, optimizing processing conditions for all types of samples is extremely important, and methods validated for one type of sample cannot automatically be adapted for other types [148]. Besides that, it is important to keep track of and consider conditions such as the length of time between sampling, freezing, and thawing that can affect the microbial community during the experiment [148].

### 5.3. DNA/RNA Extraction

The DNA/RNA extraction methodology can influence downstream sequencing data. Bead-based and column-based are the two main nucleic acid extraction approaches. For infective disease diagnosis, nucleic acid extraction methods should be selected based on their advantages and limitations, as well as the specific diagnostic requirements. Compared with conventional liquid-phase extraction methods, the bead-based extraction method is more suitable for mNGS in the context of infectious disease diagnosis because of the higher yield, purity, and reproducibility. According to McEvoy et al. (2020), the use of solid-phase reversible immobilization beads (SPRI) in DNA purification is favored by many protocols because of its ease of use, cost-effectiveness, size selectability, and automation capability [149]. Among the many methods of extracting nucleic acids from host samples, bead-based extraction was shown to be the best suited for depletion of host nucleic acids for diagnostic metagenomics of infectious diseases [150]. In diagnostic settings, column-based methods are demonstrated to be versatile and effective in detecting pathogens such as herpes simplex virus and parvovirus B19 [151,152]. However, column-based methods may result in DNA loss and lower recovery rates, highlighting the superiority of bead-based extraction for mNGS in capturing the comprehensive and high-quality metagenomic profile essential for accurate infectious disease diagnosis [153].

As mentioned before, each sample type should have unique extraction methods according to the guidelines from the FDA. Due to the differences in the composition of cell walls among plants, animals, bacteria, and fungi, the extraction of DNA and RNA for mNGS requires distinct and tailored methods. Plants with rigid cellulose-based cell walls may require methods like cetyltrimethylammonium bromide (CTAB) or using a plant-specific kit, such as the DNeasy Plant Mini Kit from Qiagen [154], while animals lacking cell walls but containing other cellular components may benefit from phenol-chloroform extraction and a column-based purification kit, such as the DNeasy Blood & Tissue Kit from Qiagen. Bacteria, with unique cell wall structures in Gram-positive and Gram-negative species, may necessitate enzymatic lysis or bead beating methods to disrupt the cell wall and membrane, in order to extract DNA and RNA [155]. Fungi with chitin-based cell walls may require enzymatic digestion for efficient nucleic acid extraction [156].

The choice of lysis approaches is also the main factor to affect the downstream sequencing result. For example, some cells are less susceptible to lysis techniques, resulting in the under-representation of DNA from these organisms [157–160]. Therefore, the extraction method must be effective for different microbial species to prevent easy-to-lyse bacteria from dominating the sequencing results. A variety of lysis approaches can be used to extract DNA, including mechanical and chemical methods. Using mechanical methods, such as bead beating, can result in DNA fragmentation and loss during library preparation methods involving fragment size selection. Hence, it is necessary to use the appropriate extraction technique, to ensure that a sample is representative of all cells present with high

quality and sufficient quantity. On the other hand, contamination occurring during the sample extraction and library preparation also can affect the quality of the downstream sequencing data. There are several sources of contamination, including the environment, personnel, machines, as well as laboratory reagents [161]. It is particularly problematic for samples with low biomass, as there is less signal to compete with low contamination levels [91]. As a precaution, ultraclean reagents are recommended and no template sequencing controls are included [91]. In RNA extraction, to get better yields from *in vitro* transcription reactions, degradation during RNA purification protocols, and lower yields from RNA extraction, decontamination with the RNase decontamination solution is crucial. It is also possible to obtain RNA from total nucleic acid by treating it with DNase. Carrier RNA should not be used for RNA mNGS, since it will be sequenced along with the sample RNA and affect the viral sequencing depth [162].

#### 5.4. Fragmentation

Before library preparation, some sequencing methods, such as short-read sequencing, require fragmenting the template DNA, cDNA, or RNA. As a result of fragmentation, DNA or RNA can be made into a size range suitable for the particular method and sequencing platform. It can be either mechanical or enzymatic fragmentation. Long RNA fragments are typically fragmented chemically. It is important to consider how fragmentation methods impact library coverage, to avoid GC biases introduced by specific fragmentation methods. Consideration should also be given to the amount of starting material and the potential sample loss associated with each method. In mNGS, fragment size can affect sequencing depth and the ability to assemble complete genomes from data [163]. For instance, small microbial genome size and limited fragment size necessitate a higher number of fragments for metagenomic sequencing [163]. Nevertheless, mNGS presents a substantial challenge due to the complexity of the data, which may consist of sequence fragments from many different genomes [164]. Due to the unassembled nature of most mNGS, conventional methods of gene discovery are not relevant [165]. To address this issue, metagenomic approaches involve the sequencing of randomly sheared fragments, known as shotgun sequencing, to provide a potentially more accurate characterization of microbial diversity [166].

#### 5.5. Library Preparation

The choice of library preparation method can significantly influence downstream sequencing quality issues, and choosing the right library preparation methods for different sample inputs is crucial for ensuring robust data quality across samples in different preservation conditions [128]. Sequencing bias in library preparation kits is important in mNGS, which involves a wide variety of microbes with varying GC contents [135,167]. Additionally, some studies have also demonstrated that library preparation procedure and sample characteristics influence metagenomic profile accuracy, which is important in choosing the best preparation method for metagenomic libraries based on sample characteristics, including community types, compositions, and DNA amounts [168,169]. According to Wang et al. (2022), different sequencing library preparation methods were evaluated to get efficient and high-fidelity metagenomic data from samples with low microbial biomass, emphasizing how library preparation is crucial for overcoming challenges associated with low biomass samples [170]. Similarly, Jones et al. (2015) found significant differences in taxonomy between next-generation sequencing library preparations, highlighting the significance of library preparation methodology for predicting genomic and functional outcomes [171].

#### 5.6. Sequencing Platform

When selecting a sequencing platform for integrating assays in mNGS, it is crucial to consider various factors, such as cost, the amount of sample input, and required concentration and fragmentation size ranges. The sequencing depth required often influences the choice of platform, as different platforms come with distinct advantages and

challenges [172,173]. Third-generation sequencing platforms, such as Oxford Nanopore Technologies (ONT), have demonstrated potential in offering precise species resolution compared to Illumina sequencing [174,175]. Parameters like read length, sequencing coverage, and sequencing errors play a significant role in the efficiency and accuracy of metagenomic sequence assembly [42,176]. Therefore, a thoughtful evaluation of the sequencing platform is necessary to ensure optimal performance and dependable outcomes in metagenomic studies.

#### Bioinformatics Analysis for Metagenomics

For metagenomic data analysis, it is often difficult to determine which sequences belong to a given pathogen in mNGS datasets, as the majority of sequences come from the host, and only a very small fraction of sequences come from the pathogen. Therefore, according to ISO24420:2023 Clause 5.2.3, it is recommended that host reads from all organisms be removed by mapping them to a genome reference such as UniRef or the Unified Human Gastrointestinal Protein Catalog [42,128]. mNGS involves a series of bioinformatics steps to analyze raw data effectively, including preprocessing of sequencing reads, de novo metagenome assembly, genome binning, taxonomic and functional analysis of genomes, fragment recruitment to reference genomes, and metagenomic assembly and analysis (Table 4). In dealing with metagenomic datasets, there are three main approaches, including marker gene analysis, binning, and assembly [115,165].

**Table 4.** Software tools involved in bioinformatics analysis for metagenomics.

Bioinformatic Step	Software	Function	Reference
Preprocessing of Sequencing Reads	FastQC (v0.12.0)	Assesses the quality of raw sequencing data	
	Trimmomatic (v0.4)	Trimming and filtering reads to eliminate low-quality bases and adapter sequences	[177]
	Cutadapt (v3.4)	Efficiently removes adapter sequence	[178]
	DUST (v0.9)	De-replication of reads	
	QIIME (v2023.2)	Noise removal	
Fragment Recruitment to Reference Genomes	Bowtie2 (v2.5.3)	Mapping preprocessed reads to reference genomes, contamination removal	[179]
	BWA (v0.7.12)	Mapping low-divergent sequences against a large reference genome	
	BWA-SW (v0.7.12) and BWA-MEM (v0.7.12)	Mapping longer sequences (70 bp to 1 Mbp), share similar features such as long-read support and split alignment	[180]
	SAMtools (v1.19.2)	Manipulating and analyzing sequence alignment data, crucial for post-processing and downstream analysis of metagenomic data	[181, 182]
De Novo Metagenome Assembly	MEGAHIT (v1.2.9)	Assembly of large and intricate metagenomic datasets using succinct de Bruijn graphs, providing a single-node solution for complex assemblies	[183, 184]
	IDBA-UD (v1.1.3)	A specialized de Bruijn graph-based assembler designed for metagenomic sequencing data, aiding in the reconstruction of microbial genomes	[185]
	MetaSPAdes (v3.13.0)	Advanced metagenomic assemblers that integrate information from multiple samples to improve accuracy and congruency	[186]
	QUAST (v5.0.2)	Evaluates genome/metagenome assemblies	[187]
Genome Binning	MaxBin (v2.0)	Binning tool that clusters contigs based on expectation-maximization algorithms, facilitating metagenomic data organization	[188]
	CONCOCT (v1.1.0)	The recovery of metagenome-assembled genomes in situations where the reference genome for a species of interest within a metagenome is unknown	[189]
	MetaBAT (v2.15)	Binning metagenomic contigs into genome bins based on sequence composition and abundance	[189]

Table 4. Cont.

Bioinformatic Step	Software	Function	Reference
Taxonomic and Functional Analysis of Genomes	Kraken 2 (v2.0.8 beta)	A highly accurate classifier for taxonomic sequences that rapidly assigns taxonomic labels to metagenomic sequences, allowing precise taxonomic analysis	[190]
	HUMAnN 3 (v3)	The tool provides insights into the functional potential of microbial populations based on metagenomic sequencing data	[191]
	MetaPhlAn 4 (v4)	This tool aids in taxonomic analysis of metagenomic shotgun sequencing data by profiling microbial communities	[192]
Metagenomic Assembly and Analysis	Anvi'o (v2.1.0)	Analyze and visualize complex metagenomic datasets using an interactive platform for metagenomic analysis and visualization	[193]
	MEGAN-LR (v6.19.1)	A long-read version of MEGAN for taxonomic analysis and functional annotation of metagenomic data generated from long-read sequencing technologies	[194]

The marker gene analysis compares each sequence with a database of taxonomically and phylogenetically informative sequences called marker genes. The sequences are then taxonomically annotated after the similarities are assessed [165]. According to ISO24420:2003, for obtaining a higher-level taxonomy profile (e.g., species, genus, order, or phylum), including metagenomic linkage groups (MLGs), metagenomic clusters (MGCs), and metagenomic species (MGS), the best taxonomy profile method should be selected based on the data and the application needs [128]. The concept of binning refers to the clustering of sequences according to taxonomic groups, such as species, genus, or higher taxa [165]. Assembly is to combine the small sequences from your sample to make longer sequences represent genomes [165]. In the absence of a reference genome dataset, such as soil or ocean samples, sequence assembly should be used [128]. In de novo assemblies, contigs or scaffolds are derived without any reference from sequence fragments [128]. The choice of sequence assembly software should take into account the extent to which accuracy, contig size, input data type, and computational resources are important [128]. Therefore, different profiles, such as phylogenetics/taxonomic abundance and contigs/genome coverage, are able to provide valuable indications for function annotation and taxonomic classification.

#### 5.7. The Choice of Reference Database

Additionally, one of the limitations of metagenomics is the quality and availability of reference databases. The choice of a reference database has a significant impact on taxonomic classification results [195]. Reference databases should contain all relevant genome sequences covering the entire genetic diversity of organisms and ensure that no artificial, low-quality, or incorrectly named genome sequences are included [196]. The incorrect assignment of ambiguously mapped/aligned reads or k-mers in poorly curated databases may lead to false positive results [196]. False negative results can be caused by insufficient databases that do not include newly discovered viral strains or uncommon strains [197]. In pathogen detection, compressing a database by removing duplicate sequences will reduce performance, but it is an effective way of saving storage space [198]. Generally, larger databases allow for better sub-typing/classification at the isolate level [196].

The National Center for Biotechnology Information (NCBI) is a popular database for metagenomic analysis [199]. However, a biased collection of isolated viruses may be present in the NCBI database [200]. Hence, researchers have manually identified metagenomic viral contigs to supplement the viral protein families found in the NCBI [200]. Through this approach, viral genome diversity can be expanded and viral communities can be better represented in metagenomic studies. The FDA-ARGOS database and the FDA Reference Viral Database (RVDB) are examples of databases that have been developed for diagnostic use and regulatory science [201], and which provide quality-controlled reference genomes for various pathogens. For instance, by using the MEGARes V3.0 database, antimicrobial resistance genes can be identified from mNGS data [118], while the Microba Commu-

nity Profiler can be used to identify taxonomic groups from metagenomic datasets [202]. According to ISO24420:2023, taxonomy profiling should be conducted using reference datasets, such as the RefSeq complete genomes (RefSeq CG) for microbes, and the BLAST database for high-quality nucleotide and protein sequences [42]. The other databases, such as ICTV Virus Metadata Resource (VMR) (VMR\_MSL38\_v2, <https://ictv.global/vmr> (accessed on 20 January 2024), Reference Viral DataBase (RVDB) (v28.0, 22 November 2023), Reference Viral DataBase (RVDB, protein version 28, <https://rvdb-prot.pasteur.fr/> (accessed on 20 January 2024), SIB Viral reference sequences, UniProt Virus proteomes, VirMet (v1.2), Virosaurus, and Virus Pathogen Resource (ViPR), are available for pathogen/viral mNGS [196]. There is also a list of simulated datasets and synthetic datasets publicly available for analysis of this type [203]. However, there may be differences between these databases when it comes to availability and coverage, and it is important to standardize data quality and representation [172]. Since reference genome databases for microorganisms are constantly updated, laboratories must pay attention to the exact versions being used as well as possible mis-annotations in reference databases [30]. A metagenomic profile can be annotated at various levels according to the reference annotation, e.g., species, genus, or higher [42].

### 5.8. Reporting

Several key findings should be reported to doctors as part of the metagenomic analysis. Firstly, the methods, limitations, and quality, including the depth of coverage of the assay, must be reported. Secondly, it is essential to report whether specific pathogens are present in the sample since this information can be used to guide treatment decisions [25]. Additionally, comparing abundances of microorganisms can help assess microbiome composition and dysbiosis [204]. Furthermore, antibiotic resistance genes and virulence factors should be reported as part of the analysis, as this information may inform treatment strategies [205]. A functional analysis should also include information from the microbiome, such as metabolic pathways or genes that play a key role in the microbiome [206]. Finally, clinical interpretation of results should be conveyed concisely and clearly, which will enable accurate diagnosis, guide treatment decisions, optimize antimicrobial therapy, and lead to improved patient outcomes [206]. For instance, the results of mNGS can give patients and clinicians confidence that unnecessary empirical treatments can be stopped and can tell them when coinfections should be ruled out, and when infectious syndromes should be diagnosed [25]. Furthermore, it may be possible to expedite or defer the escalation of empirical therapy based on mNGS results, resulting in a more timely and effective treatment for patients [207]. Through mNGS, pathogens and antimicrobial resistance determinants can also be identified. This helps tailor therapy to the specific infectious agent and potentially reduces antimicrobial resistance selection [208]. The presence of viruses of unknown pathogenicity or seldom-detected viruses may not have caused a specific disease in the past, but at a later stage could lead to a specific syndrome, which is why it is recommended to have these viruses reported [196]. An interpretation of an unknown metagenomic finding can be discussed subsequently with the clinician or noted in the report as well.

## 6. Assay Validation

According to ISO2037-1:2022 Clause 6, before accumulating validation data, a validation protocol for intended use shall be developed, implemented, and documented [43]. Through the validation process, potential failures should be identified systematically. It is also important to validate assays using samples of the type intended for use in the test, so that the test results are representative of the larger sample population [43]. Validation can be carried out with known positive and spiked specimens throughout the workflow, including pathogens that are exclusive to a particular region and those that are common colonizers, as well as microorganisms that are found as contaminants in the environment at high levels [209]. In each matrix, the limits of detection (LODs) for reference microorganisms should be determined [209]. It is important to keep in mind that sequencing depth can



affect the LOD; therefore, it is necessary to provide similar sequencing depths for all samples, which can fall into an acceptance range of sequencing depths [209]. Through repeated measurements of samples with low concentrations of analytes, the limit of quantity (LOQ) can also be used to evaluate the sensitivity and reliability of an analytical method. Besides that, for tNGS, gene- and disease-specific aspects must be included in test validations for disease-specific targeted gene panels. To ensure that the most common causes of disease are detectable, common pathogenic variants must be included in the validation set, even though sequence-specific context can affect its detection [210]. Additionally, negative controls are essential to eliminate possible contamination, which can occur at any stage from sampling to sequencing. Sample blanks, nucleic acid extraction blanks, and/or no-template controls can be used as negative controls.

Considering the complexity of NGS, thorough investigation and root cause analysis are essential for ensuring assay quality, safety, and efficacy; therefore, performance validation and verification are important steps in bioinformatics analysis for mNGS. Through error-based design and optimization, potential failure points can be identified, and the level of validation and quality control can be determined. As part of “dry lab” validation, validation of the metagenomic bioinformatics software pipeline is essential. The pipeline should be validated before being used for analytical purposes, including all tools, code, environment, and network connections [42]. The in-house bioinformatics pipelines were compared with popular metagenomics pipelines using simulated *in silico* datasets and published raw patient datasets [211]. On the other hand, wet lab validation ensures the accuracy and reliability of mNGS results as it involves experimentally verifying the sequencing results with various laboratory techniques and reagents based on ISO15189 and ISO20397 [46,211]. In both dry and wet labs, mNGS validation should be based on the 5M1E methodology to enhance problem-solving capabilities, optimize procedures, and maintain high-quality assurance standards.

### 6.1. Reference Materials

The problem is exacerbated by the lack of reference standards in many clinical scenarios [25]. For instance, the FDA advises that when comparing results with non-reference standards, specificity and sensitivity should be expressed as negative and positive percent agreement, respectively, according to their statistical guidance on reporting results from studies evaluating diagnostic tests [212]. If no virus-positive residual clinical material is available, then mock samples can be manufactured using reference material or purchased commercially [213]. For instance, the National Institute for Standards and Technology (NIST) is developing whole-genome microbial reference standards [43]. Through genome assembly, base-level analysis, genomic purity, and genomic stability evaluations, four clinically important microbial species are being characterized by NIST [43]. To validate a microbial NGS-based diagnostic test, these reference materials (RM 8375, a 4 bacteria panel, and RM 8376, a 19 bacteria + 1 human panel) can be used in appropriate steps [43], or the reference controls also can be obtained from UK-NIBSC, the European virus archive, or the ATCC [37].

As part of accuracy studies, the choice of samples is very important. A reference control material should contain multiple viruses that have different relative loads so that specificity can be determined, or separately prepared mixtures of viral strains with different viral concentrations so that sensitivity can be assessed [37]. An evaluation of a sample analyzed with a previously analyzed or another platform, such as Sanger sequencing, should be conducted as part of mNGS validation [210]. Typically, residual samples are spiked at levels exceeding the sensitivity of the test (e.g., 10 times the LOD) after evaluation [48]. However, it may not be practical to establish clinically and analytically relevant LODs for all potential targeted organisms due to the complexity of metagenomic tests [48]. Representative and positive patient specimens for the relevant specimen type can be used as spiking sources [48,209]. By doing this, one can measure background matrix effects on a larger number of independent specimens than can be done using test-positive

specimens [48]. In this step, the test needs to demonstrate its analytical performance for different microorganisms so that maximum confidence can be placed in its ability to distinguish targeted microorganisms from common microorganisms.

### 6.2. Proficiency Testing and External Quality Assessment (EQA)

To ensure the consistency and reliability of the sequencing process, quality control measures are implemented. For example, it is crucial to assess sample quality before sequencing, including DNA/RNA integrity, concentration, and potential contaminants. During library preparation, the right protocol and strategies need to be selected to minimize bias and ensure representative sequencing of the microbial community [31]. Besides that, it is important to participate in EQA programs which compare the laboratory's performance with other laboratories to ensure accuracy. The following qualitative characteristics may be addressed by EQAs of mNGS methods for viral pathogen detection: (i) accurate identification of pathogen species, (ii) quantitative parameters, and (iii) logistic performance (turnaround time) [37]. In addition to real sequencing data, the College of American Pathologists recommends that proficiency testing be combined with silico-generated sequences or modified sequences [48]. This enables the examination of algorithms and sequence databases in greater depth. In silico specimen generation and critical data analysis steps can be performed with fewer resources because a wide variety of specimens can be generated [48,209]. In some cases, clinical samples were used to verify the conclusions drawn from in silico and contrived samples [209]. During data analysis, a minimum threshold should be established based on the diagnostic approach [210]. To determine data quality, additional metrics might include the percentage of reads aligned to the human genome, the percentage of unique reads (before duplicates are removed), the proportion of bases matching the target sequences, the uniformity of coverage, and the proportion of bases uncovered [210].

### 6.3. Enhancement of Data Analysis

Typically, NGS platforms generate complex data that need to be handled and processed in multiple stages. Due to the use of separate tools for data analysis, independent of wet laboratory procedures, and the likelihood that these tools may be customized in the laboratory, a separate validation process for the analytical pipeline should be conducted during the initial test development stage. After that, the pre-validated pipeline can be used to validate both the wet laboratory and the analytical pipeline within the end-to-end test validation. The laboratory should also document any validation data provided by vendors when using commercially developed software, but an independent validation should also be performed [210]. Therefore, it is necessary to determine the parameters and thresholds for determining whether the overall sequencing run is successful, such as the number of reads aligned, the number of duplicates, the average coverage depth, and the range of insert sizes.

Meanwhile, identifying the threshold for coverage is very important to ensure that sufficient coverage and allelic fraction are achieved for variant calling and the sensitivity and specificity of analytical tests. The success of mNGS is influenced by various factors, including genetic diversity, sequencing depth, and coverage. Metagenomic contigs are normally categorized at the species level based on their sequencing depth and coverage [214]. Additionally, algorithms for estimating coverage in metagenomic datasets can help inform decision-making about which assembling approach to use, especially with uneven distributions of coverage depth [215]. The extent of coverage is important for recovering microbiome information, and it directly affects metagenomic processing fidelity [216]. Furthermore, metagenomics coverage theory emphasizes the importance of adequate coverage for meaningful results, defining what depth of sequencing is needed within an experiment [217]. Further, it is important to understand the microbial ecology and gene diversity within diverse environments for the impact of sequencing depth on the characterization of the microbiome and resistome [218]. Metagenomic datasets can be

interpreted in a straightforward manner based on robust statistical frameworks, which can be used to assess the completeness of mNGS by determining how much of the genome is covered by sequencing reads [219]. It is also crucial to estimate the metagenome coverage to ensure high sensitivity in the detection of antimicrobial resistance genes or other functional elements [220]. As mentioned before, metagenomic analyses are directly impacted by sequencing coverage depth. As a guideline for best practices in shotgun metagenomics, 1 Gb of sequencing depth per sample should be used [221]. Nevertheless, mNGS coverage requirements can vary depending on the complexity of the microbiota and the specific research objectives. To successfully sequence metagenomic samples, a minimum coverage depth is needed to identify minor variations within viral quasispecies [8]. Additionally, a 15X genome coverage limit has been determined for an assembly-based method of detecting antimicrobial resistance genes [220], which highlights the importance of detecting specific genetic elements within metagenomes with adequate coverage.

#### 6.4. Performance Characteristics

During the test validation, performance characteristics such as precision reproducibility and repeatability, analytical specificity and false positive rate, analytical sensitivity and false negative rate, and LODs should be determined to provide enough evidence on the accuracy of the test. The final analytical parameters must reflect the entire testing process.

##### 6.4.1. Precision Reproducibility and Repeatability

Various instruments, reagents, and techniques can introduce random imprecision at each step in multistep and complicated procedures. To minimize variation, instruments, reagents, and personnel must be qualified and verified. Assay validation should quantify variations that can occur. Because there are many sources of variation, it is not practical to evaluate each source separately. To ensure quality, it is recommended to assess a minimum of 3 to 20 samples with known concentration or series dilution across all steps and over a long period, including all instruments, testing personnel, and multiple lots of reagents [124,209]. Quantitative precision testing should be performed both within runs (repeatability) and within laboratories (reproducibility) by using defined LOD samples. Before acquiring validation data, acceptance criteria must be established as well. Duplicate analyses of the same sample should be performed to verify reproducibility and calculate the coefficient of variation (CV%). In addition, results from different laboratories can be compared to evaluate interlaboratory performance. A laboratory's intermediate precision is the degree to which results of tests performed on the same test items over a long period are closely aligned, considering differences in laboratory conditions such as different operators, equipment, or days [222]. Analysis and comparison of sequencing data should be carried out throughout the validation process using appropriate bioinformatics tools as well [223]. Besides that, a PCR-based approach can be used to verify the quantitative accuracy [209].

##### 6.4.2. Analytic Sensitivity and Analytical Specificity

Analytic sensitivity refers to how accurately the mNGS assay can detect and identify target microorganisms [48]. The analytical sensitivity of mNGS refers to the method's ability to detect and identify pathogens accurately, directly impacting its reliability. False negative rates are the rate at which true positive results are incorrectly classified as negative results. Both parameters are critical for determining metagenomic sequencing's effectiveness in diagnosing infectious diseases. Several variables determine the sensitivity and specificity of pathogen detection, such as nucleic acid extraction efficiency, size of the pathogen genome, quality of library preparation, number of sequencing reads derived from a given specimen (coverage), the composition of samples, background noise sequences, reference sequence availability, sequencing depth, the precision of classification algorithms, and the confidence required to identify pathogens [48]. To maximize the sensitivity of mNGS assays, it is critical to extract nucleic acids efficiently and prepare libraries in a quality manner, because the number of pathogens may change the total nucleic acid yield, resulting

in a sequence library containing both patient nucleic acid as well as microbial nucleic acid. There is evidence that mNGS can enhance pathogen detection in clinical samples, especially when low pathogen titers or high nucleic acid backgrounds may yield false negative results [30]. For example, mNGS has a high false negative rate when identifying pathogens such as *Streptococcus agalactiae* in infectious native valve endocarditis [212]. Hence, it is imperative to interpret negative results carefully, especially in samples that have a high host background, since they are at higher risk of false negative results [25]. There are other protocols available for enriching pathogen nucleic acids [81,224], depleting host nucleic acids [225], or removing part of sequencing libraries [82] to increase sensitivity.

Analytic specificity in the context of mNGS refers to the ability of the sequencing method to accurately identify and distinguish specific microbial species or genetic elements within a complex sample [204], while the false positive rate represents the percentage of negative results that are incorrectly identified as positive [226]. The analytical specificity of clinical metagenomic testing can be influenced by several factors that are not taken into account in traditional validation guidelines. These include DNA fragments contaminating reagents and consumable surfaces from a variety of microorganisms, genome sequence similarity among microorganisms, incorrect reference genome sequences, and the difference between clinical isolates and reference strains. Limited analytic specificity can occur as a result of the misclassification of microorganisms or nucleic acids or the contamination of reagents, which can be attributed to algorithms used or to reference databases [48]. The identification and mitigation of problems can be assisted by extensive *in silico* validation with particular emphasis on microbes with sequence homologies with relevant pathogens [48].

#### 6.4.3. Limit of Detection

The LOD refers to the lowest concentration or abundance of pathogens or genetic material that can reliably be detected and quantified in mNGS. It is crucial to assess the accuracy and sensitivity of mNGS in identifying pathogens in clinical samples based on this parameter. For instance, Greninger et al. (2010) demonstrated that mNGS can detect pathogens near the limits of detection for RT-PCR assays even at low concentrations. This illustrates what mNGS can achieve at low concentrations, allowing pathogens that would be difficult to identify using traditional techniques to be detected [227]. By incorporating computational approaches and modeling techniques, LODs for mNGS have been refined. Serial dilutions of a clinical sample or external controls with a known, quantifiable pathogen can be analyzed to determine the LOD, or a set of calibrated internal controls can be used. For example, receiver operating characteristic (ROC) curves have been used to refine similarity searches, and a reliable, fixed bitscore value has been calculated across the sequence of the target gene to maximize sensitivity and specificity in the detection of short-gene fragments within metagenomes [228]. To determine the LOD, cut-off thresholds are defined for coverage and sequence depth that are used in decision-making in the context of sample composition, since the host nucleic acid burden can rapidly alter this LOD [37].

### 7. Future Perspectives and Challenges for Implementing mNGS in Infectious Disease Diagnosis

The future of mNGS holds a lot of promise and potential applications in a range of fields. It has the potential to provide significant new insights into the pathogenesis of new and emerging infectious diseases, as well as immunopathological phenomena relating to host–pathogen interactions [229]. In the future, deeper insights into genetic responses, mechanisms, and genetic variation may be possible due to the technology's ability to interrogate increasingly complex microbial communities [230]. However, several challenges need to be considered for successful integration (Table 5).

**Table 5.** Challenges and strategies for implementing NGS in infectious disease diagnosis.

Challenges	Strategies	Future Perspectives	References
Turnaround Time and Costs	Cost and turnaround time reduction	Development of more cost-effective NGS platforms Optimization of bioinformatics pipelines	[30,31]
	Develop universally accepted protocol or standard	Development of consensus standards through cooperation among researchers, industry stakeholders, and regulatory bodies	[34,35]
Standardization and Quality Control	Integration of automation	Integration of automation into the whole mNGS workflow	[231,232]
	Proficiency testing and reference materials	Expansion of proficiency testing programs for mNGS	[233]
	Standards and developments in bioinformatics	Availability and accessibility of rapidly evolving software to users Usage of metabarcoding and metagenomic bioinformatics	[140,234–236]
	Collaboration and open sharing	Supporting open science initiatives through funding mechanisms and academic recognition	[237]
	Technological advancements	The standardization and quality control of mNGS	[238]
Bioinformatics and Data Analysis	User-friendly tools	Development of user-friendly software and databases	[239,240]
	Data storage and privacy	Implementation of robust data privacy and security measures	[239,240]

### 7.1. Turnaround Time and Costs

NGS can be costly, requiring specialized equipment and expertise. Clinical settings with limited resources may be unable to afford sequencing and high-performance computing infrastructure [30]. In the clinical setting, providing timely results from mNGS is crucial to rapid diagnosis in clinical settings and is also challenging. Data analysis pipelines and sequencing technologies need to be improved to reduce turnaround time [25]. Future perspectives involve enhancing the cost-effectiveness of NGS platforms and optimizing bioinformatics pipelines to achieve cost reduction. For example, the RAPIDprep assay was introduced as a simple and fast protocol for the RNA mNGS of clinical samples which provides a cause-agnostic laboratory diagnosis of infection within 24 h of sample collection by sequencing ribosomal RNA-depleted total RNA [31].

### 7.2. Standardization and Quality Control

The standardization and quality control of mNGS are crucial for ensuring data accuracy and comparability. Standardizing laboratory procedures to ensure the quality and interoperability of Big Data generated by sequencing remains a crucial issue [34]. In 2011, Field et al. noted the Genomic Standards Consortium (GSC) as driving community-based standardization activities that aimed at improving the quality and quantity of contextual data associated with public collections of genomes and metagenomes [241]. The GSC has set out a minimum level of information for a metagenomic sample, emphasizing the importance of standardizing the context [242]. However, it is challenging to standardize mNGS because there are no universally accepted protocols and standards. The development of consensus standards should be the focus of future efforts through collaboration between researchers, industry stakeholders, and regulatory bodies. The process may involve the formation of international working groups or consortiums that set standards for sample preparation, library preparation, sequencing, and bioinformatics analysis [35]. Automated technologies can also increase reproducibility and reduce human-induced variability in

metagenomics sequencing workflows [243]. For instance, PacBio collaborated with an automation partner to create fully automated protocols to prepare samples for sequencing [244]. Besides that, it is important to launch proficiency testing programs for mNGS to help the standardization and implementation of mNGS [233]. With the bioinformatics field continuing to evolve, efforts should be focused on the establishment of standardized data formats, ontologies, and metadata definitions for metagenomic sequencing data. In this way, data sharing will be facilitated, and integration and meta-analysis will be improved, ultimately resulting in better reproducibility and comparability of results [141,236,243,244]. The standardization and quality control of mNGS will be enhanced by future technological advancements through collaboration and open sharing, including novel sequencing platforms, improved bioinformatics tools, and innovative quality control assays [237,238].

### 7.3. Bioinformatics and Data Analysis

Bioinformatics tools and sequencing technologies will improve the sensitivity, specificity, and speed of sequencing. Validation studies and standardization efforts will establish WGS and mNGS as an effective and reliable diagnostic method [141]. The use of WGS can improve patient outcomes and guide targeted therapies by helping to identify genetic variants associated with treatment response, disease susceptibility, and adverse drug reactions [239]. Technological advancements, such as nanopore sequencing, will enable point-of-care sequencing and expand the accessibility of WGS in a variety of clinical settings while improving its speed, accuracy, and cost-effectiveness [240]. Ultimately, mNGS and WGS are set to transform research, diagnostics, and environmental monitoring with ongoing advancements and applications across diverse fields. Therefore, the development of secure and scalable data storage solutions and user-friendly software and databases to identify pathogens, predict antimicrobial resistance, and analyze outbreaks makes these tools more accessible to healthcare professionals [236,237].

## 8. Conclusions

In conclusion, the strategies and challenges of implementing mNGS are discussed in this review. The validation under different ISO standards, as discussed in this review, is important to minimize any errors in the workflow and to ensure the accuracy and reliability of NGS in infectious disease diagnosis. As the field of clinical mNGS continues to evolve, standard protocols, bioinformatics pipelines, and reference databases are essential for addressing the challenges and facilitating the wide adoption of mNGS in clinical microbiology laboratories. In the future, the development of reference standards and guidelines will be crucial to ensuring the accuracy, reproducibility, and reliability of mNGS.

**Author Contributions:** Writing—Original Draft Preparation, C.-M.K.; Review and Editing, C.-M.K., X.M.P., H.F.T., and A.C.-S.Y.; Conceptualization, Methodology, and Supervision, S.S.M.N., S.C.C.W., A.K.-Y.Y., and A.C.-S.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by Lim Peng Suan Charitable Trust Research Grants P0046023 and P0039295.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Conflicts of Interest:** A.C.S.Y. and A.K.Y.Y. are employed by Codex Genetics Limited. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in this manuscript. No conflicts of interest were declared by any of the other authors.

## References

1. Wu, J.; Cai, K.; Feng, Q. The Application of Metagenomic Approaches in the Management of Infectious Diseases. *Trop. Med. Surg.* **2015**, *3*, 196. [\[CrossRef\]](#)
2. Yu, X.; Jiang, W.; Shi, Y.; Ye, H.; Lin, J.-H. Applications of Sequencing Technology in Clinical Microbial Infection. *J. Cell. Mol. Med.* **2019**, *23*, 7143–7150. [\[CrossRef\]](#)
3. Markin, R.S. Manifestations of Epstein-Barr Virus-Associated Disorders in Liver. *Liver Int.* **2008**, *14*, 1–13. [\[CrossRef\]](#)
4. Maiden, M.C.J.; Rensburg, M.J.J.V.; Bray, J.E.; Earle, S.G.; Ford, S.A.; Jolley, K.A.; McCarthy, N.D. MLST Revisited: The Gene-by-Gene Approach to Bacterial Genomics. *Nat. Rev. Microbiol.* **2013**, *11*, 728–736. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Hatherell, H.-A.; Colijn, C.; Stagg, H.R.; Jackson, C.; Winter, J.R.; Abubakar, I. Interpreting whole genome sequencing for investigating tuberculosis transmission: A systematic review. *BMC Med.* **2016**, *14*, 21. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Russo, T.A.; Marr, C.M. Hypervirulent klebsiella pneumoniae. *Clin. Microbiol. Rev.* **2019**, *32*, e00001-19. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Jo, Y.; Choi, H.; Kim, S.-M.; Kim, S.-L.; Lee, B.C.; Cho, W.K. The Pepper Virome: Natural Co-Infection of Diverse Viruses and Their Quasispecies. *BMC Genom.* **2017**, *18*, 453. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Naseri, N.; Petronella, N.; Ronholm, J.; Bidawid, S.; Corneau, N. Characterization of the Genomic Diversity of Norovirus in Linked Patients Using a Metagenomic Deep Sequencing Approach. *Front. Microbiol.* **2017**, *8*, 73. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Charre, C.; Regue, H.; Dény, P.; Josset, L.; Chemin, I.; Zoulim, F.; Scholtes, C. Improved hepatitis delta virus genome characterization by single molecule full-length genome sequencing combined with VIRiONT pipeline. *J. Med. Virol.* **2023**, *95*, e28634. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Wang, Y.; Xiao, X.; Chen, S.; Huang, C.; Zhou, J.; Dai, E.; Li, Y.; Liu, L.; Huang, X.; Gao, Z.; et al. The Impact of HBV Quasispecies Features on Immune Status in HBsAg+/HBsAb+ Patients With HBV Genotype C Using Next-Generation Sequencing. *Front. Immunol.* **2021**, *12*, 775461. [\[CrossRef\]](#)
11. Tonkin-Hill, G.; Martincorena, I.; Amato, R.; Lawson, A.; Gerstung, M.; Johnston, I.; Jackson, D.; Park, N.; Lensing, S.; Quail, M.A.; et al. Patterns of Within-Host Genetic Diversity in SARS-CoV-2. *eLife* **2021**, *10*, e66857. [\[CrossRef\]](#)
12. Butler, D.; Mozsary, C.; Meydan, C.; Foox, J.; Rosiene, J.; Shaiber, A.; Danko, D.; Afshinnekoo, E.; MacKay, M.; Sedlazeck, F.J.; et al. Shotgun Transcriptome, Spatial Omics, and Isothermal Profiling of SARS-CoV-2 Infection Reveals Unique Host Responses, Viral Diversification, and Drug Interactions. *Nat. Commun.* **2021**, *12*, 1660. [\[CrossRef\]](#)
13. Robinson, E.R.; Walker, T.M.; Pallen, M.J. Genomics and outbreak investigation: From sequence to consequence. *Genome Med.* **2013**, *5*, 36. [\[CrossRef\]](#)
14. Besser, J.M.; Carleton, H.A.; Trees, E.; Stroika, S.G.; Hise, K.; Wise, M.; Gerner-Smidt, P. Interpretation of Whole-Genome Sequencing for Enteric Disease Surveillance and Outbreak Investigation. *Foodborne Pathog. Dis.* **2019**, *16*, 504–512. [\[CrossRef\]](#)
15. Harris, S.R.; Cartwright, E.J.P.; Török, M.E.; Holden, M.T.G.; Brown, N.M.; Ogilvy-Stuart, A.L.; Ellington, M.J.; Quail, M.A.; Bentley, S.D.; Parkhill, J.; et al. Whole-genome sequencing for analysis of an outbreak of meticillin-resistant *Staphylococcus aureus*: A descriptive study. *Lancet Infect. Dis.* **2013**, *13*, 130–136. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Besser, J.; Carleton, H.A.; Gerner-Smidt, P.; Lindsey, R.L.; Trees, E. Next-generation sequencing technologies and their application to the study and control of bacterial infections. *Clin. Microbiol. Infect.* **2018**, *24*, 335–341. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Jackson, B.R.; Tarr, C.; Strain, E.; Jackson, K.A.; Conrad, A.; Carleton, H.; Katz, L.S.; Stroika, S.; Gould, L.H.; Mody, R.K.; et al. Implementation of Nationwide Real-time Whole-genome Sequencing to Enhance Listeriosis Outbreak Detection and Investigation. *Clin. Infect. Dis.* **2016**, *63*, 380–386. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Walker, T.M.; Ip, C.L.C.; Harrell, R.H.; Evans, J.D.; Kapatai, G.; Dediccoat, M.; Eyre, D.W.; Wilson, D.J.; Hawkey, P.M.; Crook, D.W.; et al. Whole-Genome Sequencing to Delineate Mycobacterium Tuberculosis Outbreaks: A Retrospective Observational Study. *Lancet Infect. Dis.* **2013**, *13*, 137–146. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Ladner, J.T.; Grubaugh, N.D.; Pybus, O.G.; Andersen, K.G. Precision epidemiology for infectious disease control. *Nat. Med.* **2019**, *25*, 206–211. [\[CrossRef\]](#)
20. James, S.E.; Ngcapu, S.; Kanzi, A.M.; Tegally, H.; Fonseca, V.; Giandhari, J.; Wilkinson, E.; Chimukangara, B.; Pillay, S.; Singh, L.; et al. High Resolution analysis of Transmission Dynamics of Sars-Cov-2 in Two Major Hospital Outbreaks in South Africa Leveraging Intra-host Diversity. *medRxiv* **2020**, *7*, veab041. [\[CrossRef\]](#)
21. Gu, W.; Miller, S.; Chiu, C.Y. Clinical Metagenomic Next-Generation Sequencing for Pathogen Detection. *Annu. Rev. Pathol. Mech. Dis.* **2019**, *14*, 319–338. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Campbell, F.; Strang, C.; Ferguson, N.; Cori, A.; Jombart, T. When are pathogen genome sequences informative of transmission events? *PLoS Pathog.* **2018**, *14*, e1006885. [\[CrossRef\]](#)
23. Hirabayashi, A.; Kajihara, T.; Yahara, K.; Shibayama, K.; Sugai, M. Impact of the COVID-19 Pandemic on the Surveillance of Antimicrobial Resistance. *J. Hosp. Infect.* **2021**, *117*, 147–156. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Veepanattu, P.; Singh, S.; Mendelson, M.; Nampoothiri, V.; Edathadatil, F.; Surendran, S.; Bonaconsa, C.; Mbamalu, O.; Ahuja, S.; Birgand, G.; et al. Building Resilient and Responsive Research Collaborations to Tackle Antimicrobial Resistance—Lessons Learnt From India, South Africa, and UK. *Int. J. Infect. Dis.* **2020**, *100*, 278–282. [\[CrossRef\]](#)
25. Wilson, M.R.; Sample, H.; Zorn, K.C.; Arevalo, S.; Yu, G.; Neuhaus, J.; Federman, S.; Stryke, D.; Briggs, B.; Langelier, C.; et al. Clinical Metagenomic Sequencing for Diagnosis of Meningitis and Encephalitis. *N. Engl. J. Med.* **2019**, *380*, 2327–2340. [\[CrossRef\]](#) [\[PubMed\]](#)

26. Fan, X.; Wang, Q.; Li, P.; Ai, B.; Song, Y.; Peng, Q.; Wang, H. The Diagnostic Value of Metagenomic Next-Generation Sequencing in *Angiostrongylus Cantonensis* Encephalitis/Meningitis. *J. Behav. Brain Sci.* **2021**, *11*, 216–226. [[CrossRef](#)]
27. Wang, S.; Chen, Y.; Wang, D.; Wu, Y.; Zhao, D.; Zhang, J.; Xie, H.; Gong, Y.; Sun, R.; Nie, X.; et al. The Feasibility of Metagenomic Next-Generation Sequencing to Identify Pathogens Causing Tuberculous Meningitis in Cerebrospinal Fluid. *Front. Microbiol.* **2019**, *10*, 1993. [[CrossRef](#)]
28. Liang, X.; Wang, Q.; Liu, J.; Ma, J.; Zhang, Y.; Wang, M.; Yu, Y.; Wang, L. Coinfection of SARS-CoV-2 and Influenza A (H<sub>3</sub>N<sub>2</sub>) Detected in Bronchoalveolar Lavage Fluid of a Patient With Long COVID Using Metagenomic Next-generation Sequencing: A Case Report. *Front. Cell. Infect. Microbiol.* **2023**, *13*, 1224794. [[CrossRef](#)]
29. Xie, Y.; Dai, B.; Zhou, X.; Liu, H.; Wu, W.; Yu, F.R.; Zhu, B. Diagnostic Value of Metagenomic Next-Generation Sequencing for Multi-Pathogenic Pneumonia in HIV-Infected Patients. *Infect. Drug Resist.* **2023**, *16*, 607–618. [[CrossRef](#)]
30. Chiu, C.Y.; Miller, S.A. Clinical Metagenomics. *Nat. Rev. Genet.* **2019**, *20*, 341–355. [[CrossRef](#)]
31. Tulloch, R.L.; Kim, K.; Sikazwe, C.; Michie, A.; Burrell, R.A.; Holmes, E.C.; Dwyer, D.E.; Britton, P.N.; Kok, J.; Eden, J.S. RAPIDprep: A Simple, Fast Protocol for RNA Metagenomic Sequencing of Clinical Samples. *Viruses* **2023**, *15*, 1006. [[CrossRef](#)]
32. Hoopen, P.T.; Finn, R.D.; Bongo, L.A.; Corre, E.; Fosso, B.; Meyer, F.; Mitchell, A.; Pelletier, E.; Pesole, G.; Santamaria, M.; et al. The Metagenomic Data Life-Cycle: Standards and Best Practices. *GigaScience* **2017**, *6*, gix047. [[CrossRef](#)]
33. Sala, C.; Mordhorst, H.; Grützke, J.; Brinkmann, A.; Petersen, T.N.; Poulsen, C.S.; Cotter, P.D.; Crispie, F.; Ellis, R.J.; Castellani, G.; et al. Metagenomics-Based Proficiency Test of Smoked Salmon Spiked With a Mock Community. *Microorganisms* **2020**, *8*, 1861. [[CrossRef](#)]
34. Brumfield, K.D.; Huq, A.; Colwell, R.R.; Olds, J.L.; Leddy, M.B. Microbial Resolution of Whole Genome Shotgun and 16S Amplicon Metagenomic Sequencing Using Publicly Available NEON Data. *PLoS ONE* **2020**, *15*, e0228899. [[CrossRef](#)]
35. Imanian, B.; Donaghy, J.; Jackson, T.; Gummalla, S.; Ganesan, B.; Baker, R.C.; Henderson, M.; Butler, E.K.; Hong, Y.; Ring, B.; et al. The power, potential, benefits, and challenges of implementing high-throughput sequencing in food safety systems. *npj Sci. Food* **2022**, *6*, 35. [[CrossRef](#)] [[PubMed](#)]
36. Gvender, K.; Eyre, D.W. Benchmarking Taxonomic Classifiers with Illumina and Nanopore Sequence Data for Clinical Metagenomic Diagnostic Applications. *Microb. Genom.* **2022**, *8*, 000886. [[CrossRef](#)] [[PubMed](#)]
37. López-Labrador, F.X.; Brown, J.R.; Fischer, N.; Harvala, H.; Van Boheemen, S.; Cinek, O.; Sayiner, A.; Madsen, T.V.; Auvinen, E.; Kufner, V.; et al. Recommendations for the introduction of metagenomic high-throughput sequencing in clinical virology, part I: Wet lab procedure. *J. Clin. Virol.* **2021**, *134*, 104691. [[CrossRef](#)] [[PubMed](#)]
38. Parker, K.; Wood, H.; Russell, J.A.; Yarmosh, D.; Shteyman, A.; Bagnoli, J.; Knight, B.; Aspinwall, J.R.; Jacobs, J.L.; Werking, K.; et al. Development and Optimization of an Unbiased, Metagenomics-Based Pathogen Detection Workflow for Infectious Disease and Biosurveillance Applications. *Trop. Med. Infect. Dis.* **2023**, *8*, 121. [[CrossRef](#)] [[PubMed](#)]
39. Damme, R.V.; Hölzer, M.; Viehweger, A.; Müller, B.; Bongcam-Rudloff, E.; Brandt, C. Metagenomics Workflow for Hybrid Assembly, Differential Coverage Binning, Metatranscriptomics and Pathway Analysis (MUFFIN). *PLoS Comput. Biol.* **2021**, *17*, e1008716. [[CrossRef](#)]
40. Churchward, B.; Millet, M.; Bihouée, A.; Fertin, G.; Chaffron, S. MAGNETO: An Automated Workflow for Genome-Resolved Metagenomics. *mSystems* **2022**, *7*, e0043222. [[CrossRef](#)] [[PubMed](#)]
41. Werbin, Z.R.; Hackos, B.; Dietze, M.C.; Bhatnagar, J. The National Ecological Observatory Network's Soil Metagenomes: Assembly and Basic Analysis. *F1000Research* **2021**, *10*, 299. [[CrossRef](#)] [[PubMed](#)]
42. ISO/TS 24420:2023; Parallel DNA Sequencing—General Requirements for Data Processing of Shotgun Metagenomic Sequences. International Organization for Standardization: Geneva, Switzerland, 2023.
43. ISO 20397-1:2021; Biotechnology—Massively Parallel Sequencing—Part 1: Nucleic Acid and Library Preparation. International Organization for Standardization: Geneva, Switzerland, 2022.
44. ISO 20397-2:2021; Biotechnology Massively Parallel Sequencing—Part 2: Quality Evaluation of Sequencing Data. International Organization for Standardization: Geneva, Switzerland, 2022.
45. Burd, E.M. Validation of Laboratory-Developed Molecular Assays for Infectious Diseases. *Clin. Microbiol. Rev.* **2010**, *23*, 550–576. [[CrossRef](#)] [[PubMed](#)]
46. Maschietto, C.; Otto, G.; Rouzè, P.; Debortoli, N.; Bihin, B.; Nyinkeu, L.; Denis, O.; Huang, T.-D.; Mullier, F.; Bogaerts, P.; et al. Minimal Requirements for ISO15189 Validation and Accreditation of Three Next Generation Sequencing Procedures for SARS-CoV-2 Surveillance in Clinical Setting. *Sci. Rep.* **2023**, *13*, 6934. [[CrossRef](#)] [[PubMed](#)]
47. ISO 15189:2022; Medical Laboratories Requirements for Quality and Competence. International Organization for Standardization: Geneva, Switzerland, 2022.
48. Schlaberg, R.; Chiu, C.Y.; Miller, S.; Procop, G.W.; Weinstock, G.; the Professional Practice Committee and Committee on Laboratory Practices of the American Society for Microbiology; the Microbiology Resource Committee of the College of American Pathologists. Validation of Metagenomic Next-Generation Sequencing Tests for Universal Pathogen Detection. *Arch. Pathol. Lab. Med.* **2017**, *141*, 776–786. [[CrossRef](#)] [[PubMed](#)]
49. Infectious Disease Next Generation Sequencing Based Diagnostic Devices: Microbial Identification and Detection of Antimicrobial Resistance and Virulence Markers; Draft Guidance for Industry and Food and Drug Administration Staff. Available online: <https://www.federalregister.gov/d/2016-11237> (accessed on 13 May 2016).



50. Ojala, T.; Kankuri, E.; Kankainen, M. Understanding human health through metatranscriptomics. *Trends Mol. Med.* **2023**, *29*, 376–389. [[CrossRef](#)]
51. Zampieri, G.; Campanaro, S.; Angione, C.; Treu, L. Metatranscriptomics-guided genome-scale metabolic modeling of microbial communities. *Cell Rep. Methods* **2023**, *3*, 100383. [[CrossRef](#)]
52. Fuentes-Trillo, A.; Monzó, C.; Manzano, I.; Santiso-Bellón, C.; Andrade, J.d.S.R.d.; Gozalbo-Rovira, R.; García-García, A.-B.; Rodríguez-Díaz, J.; Chaves, F.J. Benchmarking Different Approaches for Norovirus Genome Assembly in Metagenome Samples. *BMC Genom.* **2021**, *22*, 849. [[CrossRef](#)]
53. Shi, H.; Zhou, Y.; Jia, E.; Pan, M.; Bai, Y.; Ge, Q. Bias in RNA-seq Library Preparation: Current Challenges and Solutions. *BioMed Res. Int.* **2021**, *2021*, 6647597. [[CrossRef](#)]
54. Gatcliffe, C.; Rao, A.; Brigger, M.T.; Dimmock, D.; Hansen, C.H.; Montgomery, J.; Salais, R.; Coufal, N.G.; Farnaes, L. Metagenomic Sequencing and Evaluation of the Host Response in the Pediatric Aerodigestive Population. *Pediatr. Pulmonol.* **2020**, *56*, 516–524. [[CrossRef](#)] [[PubMed](#)]
55. Kolmogorov, M.; Bickhart, D.M.; Behsaz, B.; Gurevich, A.; Rayko, M.; Shin, S.B.; Kuhn, K.L.; Yuan, J.; Polevikov, E.; Smith, T.P.L.; et al. metaFlye: Scalable Long-Read Metagenome Assembly Using Repeat Graphs. *Nat. Chem. Biol.* **2020**, *17*, 1103–1110. [[CrossRef](#)] [[PubMed](#)]
56. Nicholls, S.M.; Quick, J.; Tang, S.W.; Loman, N.J. Ultra-Deep, Long-Read Nanopore Sequencing of Mock Microbial Community Standards. *GigaScience* **2019**, *8*, giz043. [[CrossRef](#)]
57. Ong, C.T.; Ross, E.M.; Boe-Hansen, G.; Turni, C.; Hayes, B.J.; Lew-Tabor, A.E. Technical Note: Overcoming Host Contamination in Bovine Vaginal Metagenomic Samples With Nanopore Adaptive Sequencing. *J. Anim. Sci.* **2021**, *100*, skab344. [[CrossRef](#)]
58. Sim, M.; Lee, J.; Wy, S.; Park, N.; Lee, D.; Kwon, D.; Kim, J. Generation and Application of Pseudo-long Reads for Metagenome Assembly. *GigaScience* **2022**, *11*, giac044. [[CrossRef](#)]
59. Meslier, V.; Quinquis, B.; Silva, K.D.; Oñate, F.P.; Pons, N.; Roume, H.; Podar, M.; Almeida, M. Benchmarking Second and Third-Generation Sequencing Platforms for Microbial Metagenomics. *Sci. Data* **2022**, *9*, 694. [[CrossRef](#)]
60. An, N.; Wang, C.; Dou, X.; Liu, X.; Wu, J.; Chen, Y. Comparison of 16S rDNA Amplicon Sequencing With the Culture Method for Diagnosing Causative Pathogens in Bacterial Corneal Infections. *Transl. Vis. Sci. Technol.* **2022**, *11*, 29. [[CrossRef](#)]
61. Szoboszlai, M.; Schramm, L.; Pinzauti, D.; Scerri, J.; Sandionigi, A.; Biazzo, M. Nanopore Is Preferable Over Illumina for 16S Amplicon Sequencing of the Gut Microbiota When Species-Level Taxonomic Classification, Accurate Estimation of Richness, or Focus on Rare Taxa Is Required. *Microorganisms* **2023**, *11*, 804. [[CrossRef](#)] [[PubMed](#)]
62. Shay, J.; Haniford, L.S.E.; Cooper, A.; Carrillo, C.D.; Blais, B.W.; Lau, C.H.-F. Exploiting a Targeted Resistome Sequencing Approach in Assessing Antimicrobial Resistance in Retail Foods. *Environ. Microbiome* **2023**, *18*, 25. [[CrossRef](#)]
63. Siljanen, H.; Manoharan, L.; Hilts, A.S.; Bagnoud, A.; Alves, R.; Jones, C.M.; Sousa, F.L.; Hallin, S.; Biasi, C.; Schleper, C. Targeted Metagenomics Using Probe Capture Detect a Larger Diversity of Nitrogen and Methane Cycling Genes in Complex Microbial Communities Than Traditional Metagenomics. *bioRxiv* **2022**. [[CrossRef](#)]
64. Hoang, M.T.V.; Irinyi, L.; Meyer, W. Long-Read Sequencing in Fungal Identification. *Microbiol. Aust.* **2022**, *43*, 14–18. [[CrossRef](#)]
65. Pei, X.M.; Yeung, M.H.Y.; Wong, A.N.N.; Tsang, H.F.; Yu, A.C.S.; Yim, A.K.Y.; Wong, S.C.C. Targeted Sequencing Approach and Its Clinical Applications for the Molecular Diagnosis of Human Diseases. *Cells* **2023**, *12*, 493. [[CrossRef](#)] [[PubMed](#)]
66. Kim, C.; Pongpanich, M.; Porntaveetus, T. Unraveling metagenomics through long-read sequencing: A comprehensive review. *J. Transl. Med.* **2024**, *22*, 111. [[CrossRef](#)] [[PubMed](#)]
67. Benoit, G.; Raguideau, S.; James, R.; Phillippy, A.M.; Chikhi, R.; Quince, C. High-quality metagenome assembly from long accurate reads with metaMDBG. *Nat. Biotechnol.* **2024**, 1–6. [[CrossRef](#)]
68. Mastrorosa, F.K.; Miller, D.E.; Eichler, E.E. Applications of long-read sequencing to Mendelian genetics. *Genome Med.* **2023**, *15*, 42. [[CrossRef](#)]
69. Silverman, J.D.; Bloom, R.J.; Jiang, S.; Durand, H.K.; Dallow, E.; Mukherjee, S.; David, L.A. Measuring and mitigating PCR bias in microbiota datasets. *PLoS Comput. Biol.* **2021**, *17*, e1009113. [[CrossRef](#)] [[PubMed](#)]
70. Gargis, A.S.; Kalman, L.V.; Lubin, I.M. Assuring the Quality of Next-Generation Sequencing in Clinical Microbiology and Public Health Laboratories. *J. Clin. Microbiol.* **2016**, *54*, 2857–2865. [[CrossRef](#)]
71. Zhong, Y.; Xu, F.; Wu, J.; Schubert, J.; Li, M.M. Application of Next Generation Sequencing in Laboratory Medicine. *Ann. Lab. Med.* **2021**, *41*, 25–43. [[CrossRef](#)]
72. Мурашко, Л.А.; Муха, Т.И.; Humenyuk, O.; Kirilenko, V.; Novytska, N.V. The Level of Intensity of Soft Winter Wheat Varieties Infection by Fusarium Link Pathogens and Their Identification on Grain. *Plant Soil Sci.* **2022**, *13*, 35–45. [[CrossRef](#)]
73. Ambrose, M.; Aj, K.; Formenty, P.; Muyembe-Tamfum, J.J.; Aw, R.; Lloyd-Smith, J.O. Quantifying Transmission of Emerging Zoonoses: Using Mathematical Models to Maximize the Value of Surveillance Data. *bioRxiv* **2019**. [[CrossRef](#)]
74. Jones, K.E.; Patel, N.; Levy, M.J.; Storeygard, A.; Balk, D.; Gittleman, J.L.; Daszak, P. Global Trends in Emerging Infectious Diseases. *Nature* **2008**, *451*, 990–993. [[CrossRef](#)]
75. Plowright, R.K.; Parrish, C.R.; McCallum, H.; Hudson, P.J.; Ko, A.I.; Graham, A.L.; Lloyd-Smith, J.O. Pathways to Zoonotic Spillover. *Nat. Rev. Microbiol.* **2017**, *15*, 502–510. [[CrossRef](#)]
76. Rizzoli, A.; Tagliapietra, V.; Cagnacci, F.; Marini, G.; Arnoldi, D.; Rosso, F.; Rosà, R. Parasites and Wildlife in a Changing World: The Vector-Host- Pathogen Interaction as a Learning Case. *Int. J. Parasitol. Parasites Wildl.* **2019**, *9*, 394–401. [[CrossRef](#)]

77. Wu, L.; Wu, Z.C.; Тодосійчук, Т.С.; Korneva, O. Nosocomial Infections: Pathogenicity, Resistance and Novel Antimicrobials. *Innov. Biosyst. Bioeng.* **2021**, *5*, 73–78. [[CrossRef](#)]
78. Heesterbeek, J.A.P.; Anderson, R.M.; Andreasen, V.; Bansal, S.; Angelis, D.D.; Dye, C.; Eames, K.T.D.; Edmunds, W.J.; Frost, S.D.W.; Funk, S.; et al. Modeling Infectious Disease Dynamics in the Complex Landscape of Global Health. *Science* **2015**, *347*, aaa4339. [[CrossRef](#)]
79. Salipante, S.J.; Kawashima, T.; Rosenthal, C.; Hoogestraat, D.R.; Cummings, L.A.; Sengupta, D.J.; Harkins, T.T.; Cookson, B.T.; Hoffman, N.G. Performance comparison of Illumina and ion torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling. *Appl. Environ. Microbiol.* **2014**, *80*, 7583–7591. [[CrossRef](#)]
80. Weyrich, L.S.; Farrer, A.G.; Eisenhofer, R.; Arriola, L.A.; Young, J.M.; Selway, C.A.; Handsley-Davis, M.; Adler, C.; Breen, J.; Cooper, A. Laboratory Contamination Over Time During Low-biomass Sample Analysis. *Mol. Ecol. Resour.* **2019**, *19*, 982–996. [[CrossRef](#)]
81. Shi, Y.; Wang, G.; Lau, H.C.-H.; Yu, J. Metagenomic Sequencing for Microbial DNA in Human Samples: Emerging Technological Advances. *Int. J. Mol. Sci.* **2022**, *23*, 2181. [[CrossRef](#)]
82. Gu, W.; Crawford, E.D.; O'Donovan, B.; Wilson, M.R.; Chow, E.D.; Retallack, H.; DeRisi, J.L. Depletion of Abundant Sequences by Hybridization (DASH): Using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. *Genome Biol.* **2016**, *17*, 41. [[CrossRef](#)] [[PubMed](#)]
83. Ong, C.T.; Boe-Hansen, G.; Ross, E.M.; Blackall, P.J.; Hayes, B.J.; Lew-Tabor, A.E. Evaluation of Host Depletion and Extraction Methods for Shotgun Metagenomic Analysis of Bovine Vaginal Samples. *Microbiol. Spectr.* **2022**, *10*, e0041221. [[CrossRef](#)] [[PubMed](#)]
84. Gan, M.; Wu, B.; Yan, G.; Li, G.; Sun, L.; Li, G.; Zhou, W. Combined Nanopore Adaptive Sequencing and Enzyme-Based Host Depletion Efficiently Enriched Microbial Sequences and Identified Missing Respiratory Pathogens. *BMC Genom.* **2021**, *22*, 732. [[CrossRef](#)]
85. Jin, N.; Kan, C.-M.; Pei, X.M.; Cheung, W.L.; Ng, S.S.M.; Wong, H.T.; Cheng, H.Y.-L.; Leung, W.W.; Wong, Y.N.; Tsang, H.F.; et al. Cell-free circulating tumor RNAs in plasma as the potential prognostic biomarkers in colorectal cancer. *Front. Oncol.* **2023**, *13*, 1134445. [[CrossRef](#)] [[PubMed](#)]
86. Aja-Macaya, P.; Rumbo-Feal, S.; Poza, M.; Cañizares, A.; Vallejo, J.A.; Bou, G. A new and efficient enrichment method for metagenomic sequencing of Monkeypox virus. *BMC Genom.* **2023**, *24*, 29. [[CrossRef](#)]
87. Barzon, L.; Lavezzo, E.; Costanzi, G.; Franchin, E.; Toppo, S.; Palù, G. Next-Generation Sequencing Technologies in Diagnostic Virology. *J. Clin. Virol.* **2013**, *58*, 346–350. [[CrossRef](#)]
88. Gwinn, M.; MacCannell, D.; Armstrong, G.L. Next-Generation Sequencing of Infectious Pathogens. *JAMA* **2019**, *321*, 893–894. [[CrossRef](#)]
89. Poulsen, C.S.; Kaas, R.S.; Aarestrup, F.M.; Pamp, S.J. Standard Sample Storage Conditions Have an Impact on Inferred Microbiome Composition and Antimicrobial Resistance Patterns. *Microbiol. Spectr.* **2021**, *9*, e01387-21. [[CrossRef](#)]
90. Bundgaard-Nielsen, C.; Hagstrøm, S.; Sørensen, S. Interpersonal Variations in Gut Microbiota Profiles Supersedes the Effects of Differing Fecal Storage Conditions. *Sci. Rep.* **2018**, *8*, 17367. [[CrossRef](#)]
91. Salter, S.J.; Cox, M.J.; Turek, E.; Calus, S.T.; Cookson, W.; Moffatt, M.F.; Turner, P.; Parkhill, J.; Loman, N.J.; Walker, A.W. Reagent and Laboratory Contamination Can Critically Impact Sequence-Based Microbiome Analyses. *BMC Biol.* **2014**, *12*, 87. [[CrossRef](#)]
92. Marchukov, D.; Li, J.; Juillerat, P.; Misselwitz, B.; Yilmaz, B. Benchmarking Microbial DNA Enrichment Protocols From Human Intestinal Biopsies. *Front. Genet.* **2023**, *14*, 1184473. [[CrossRef](#)]
93. Bicalho, M.L.S.; Machado, V.S.; Higgins, C.H.; Lima, F.S.; Bicalho, R.C. Genetic and Functional Analysis of the Bovine Uterine Microbiota. Part I: Metritis Versus Healthy Cows. *J. Dairy Sci.* **2017**, *100*, 3850–3862. [[CrossRef](#)] [[PubMed](#)]
94. Votrubová, J.; Saskova, L.; Dalihodová, S.; Vaněk, D. DNA Extraction From Forensic Samples Using MagCore HF 16 Plus Automated Nucleic Acid Extractor—A Preliminary Study. *Forensic Sci. Int. Genet. Suppl. Ser.* **2017**, *6*, e150–e152. [[CrossRef](#)]
95. Yu, Z.; Morrison, M. Improved Extraction of PCR-quality Community DNA From Digesta and Fecal Samples. *Biotechniques* **2004**, *36*, 808–812. [[CrossRef](#)] [[PubMed](#)]
96. Oñate, F.P.; Batto, J.-M.; Juste, C.; Fadlallah, J.; Fougeroux, C.; Gouas, D.; Pons, N.; Kennedy, S.; Levenez, F.; Doré, J.; et al. Quality Control of Microbiota Metagenomics by K-Mer Analysis. *BMC Genom.* **2015**, *16*, 183. [[CrossRef](#)]
97. Riemann, K.; Adamzik, M.; Frauenrath, S.; Egensperger, R.; Schmid, K.W.; Brockmeyer, N.H.; Siffert, W. Comparison of Manual and Automated Nucleic Acid Extraction From Whole-blood Samples. *J. Clin. Lab. Anal.* **2007**, *21*, 244–248. [[CrossRef](#)]
98. Babadi, Z.K.; Narmani, A.; Ebrahimipour, G.; Wink, J. Combination and Improvement of Conventional DNA Extraction Methods in Actinobacteria to Obtain High-Quantity and High-Quality DNA. *Iran. J. Microbiol.* **2022**, *14*, 186–193. [[CrossRef](#)]
99. Bachmann, N.L.; Rockett, R.; Timms, V.J.; Sintchenko, V. Advances in Clinical Sample Preparation for Identification and Characterization of Bacterial Pathogens Using Metagenomics. *Front. Public Health* **2018**, *6*, 363. [[CrossRef](#)]
100. Zhang, L.; Chen, T.; Wang, Y.; Zhang, S.; Lv, Q.; Kong, D.; Jiang, H.; Zheng, Y.; Ren, Y.; Huang, W.; et al. Comparison Analysis of Different DNA Extraction Methods on Suitability for Long-Read Metagenomic Nanopore Sequencing. *Front. Cell. Infect. Microbiol.* **2022**, *12*, 919903. [[CrossRef](#)]
101. Bowers, R.M.; Clum, A.; Tice, H.; Lim, J.; Singh, K.P.; Ciobanu, D.; Ngan, C.Y.; Cheng, J.F.; Tringe, S.G.; Woyke, T. Impact of Library Preparation Protocols and Template Quantity on the Metagenomic Reconstruction of a Mock Microbial Community. *BMC Genom.* **2015**, *16*, 856. [[CrossRef](#)]

102. Sedláčková, T.; Repiská, G.; Celec, P.; Szemes, T.; Minárik, G. Fragmentation of DNA Affects the Accuracy of the DNA Quantitation by the Commonly Used Methods. *Biol. Proced. Online* **2013**, *15*, 5. [[CrossRef](#)] [[PubMed](#)]
103. Francesconi, A.; Kasai, M.; Harrington, S.M.; Beveridge, M.; Petraitienė, R.; Petraitis, V.; Schaefe, R.L.; Walsh, T.J. Automated and Manual Methods of DNA Extraction For *Aspergillus fumigatus* and *Rhizopus oryzae* Analyzed by Quantitative Real-Time PCR. *J. Clin. Microbiol.* **2008**, *46*, 1978–1984. [[CrossRef](#)] [[PubMed](#)]
104. Miller, D.N.; Bryant, J.; Madsen, E.L.; Ghiorse, W.C. Evaluation and Optimization of DNA Extraction and Purification Procedures for Soil and Sediment Samples. *Appl. Environ. Microbiol.* **1999**, *65*, 4715–4724. [[CrossRef](#)] [[PubMed](#)]
105. Trivedi, C.B.; Keuschnig, C.; Larose, C.; Rissi, D.V.; Mourot, R.; Bradley, J.A.; Winkel, M.; Benning, L.G. DNA/RNA Preservation in Glacial Snow and Ice Samples. *Front. Microbiol.* **2022**, *13*, 894893. [[CrossRef](#)] [[PubMed](#)]
106. Ojwang, R.A.; Adan, A.A.; Nyaboga, E.N.; Muge, E.K.; Mbatia, B.N.; Ogoyi, D.O. Optimised Germination Protocol for Jackfruit Seeds and Evaluation of Methods for Extraction of DNA Suitable for Genetic Analysis. *Afr. Crop Sci. J.* **2022**, *30*, 271–281. [[CrossRef](#)]
107. Jia, Y.; Zhao, S.; Guo, W.; Peng, L.; Zhao, F.; Wang, L.; Fan, G.; Zhu, Y.; Xu, D.; Liu, G.; et al. Sequencing introduced false positive rare taxa lead to biased microbial community diversity, assembly, and interaction interpretation in amplicon studies. *Environ. Microbiome* **2022**, *17*, 43. [[CrossRef](#)] [[PubMed](#)]
108. Larin, A.K.; Klimina, K.M.; Veselovsky, V.A.; Olekhovich, E.I.; Morozov, M.D.; Boldyreva, D.I.; Yunes, R.A.; Manolov, A.I.; Fedorov, D.E.; Pavlenko, A.V.; et al. An improved and extended dual-index multiplexed 16S rRNA sequencing for the Illumina HiSeq and MiSeq platform. *BMC Genom. Data* **2024**, *25*, 8. [[CrossRef](#)] [[PubMed](#)]
109. Gomez-Alvarez, V.; Teal, T.K.; Schmidt, T.M. Systematic Artifacts in Metagenomes from Complex Microbial Communities. *ISME J.* **2009**, *3*, 1314–1317. [[CrossRef](#)] [[PubMed](#)]
110. Hoff, K.J. The effect of sequencing errors on metagenomic gene prediction. *BMC Genom.* **2009**, *10*, 520. [[CrossRef](#)] [[PubMed](#)]
111. Jia, B.; Liu, X.; Cai, K.; Hu, Z.; Ma, L.; Wei, C. NeSSM: A Next-Generation Sequencing Simulator for Metagenomics. *PLoS ONE* **2013**, *8*, e75448. [[CrossRef](#)]
112. An, L.; Pookhao, N.; Jiang, H.; Xu, J. Statistical Approach of Functional Profiling for a Microbial Community. *PLoS ONE* **2014**, *9*, e106588. [[CrossRef](#)]
113. Walsh, A.M.; Crispie, F.; O’Sullivan, Ó.; Finnegan, L.; Claesson, M.J.; Cotter, P.D. Species Classifier Choice Is a Key Consideration When Analysing Low-Complexity Food Microbiome Data. *Microbiome* **2018**, *6*, 50. [[CrossRef](#)]
114. Zehrh, I.; Habiba, U.; Picco, M.R.; Bashir, S.H.; Rehman, U.A.; Haider, O.; Khoso, S. Metagenomics and Machine Learning-Based Precision Medicine Approaches for Autoimmune Diseases. *Preprints* **2023**, 2023040209. [[CrossRef](#)]
115. Sharpton, T.J. An introduction to the analysis of shotgun metagenomic data. *Front. Plant Sci.* **2014**, *5*, 209. [[CrossRef](#)]
116. Sedlář, K.; Kupkova, K.; Provazník, I. Bioinformatics Strategies for Taxonomy Independent Binning and Visualization of Sequences in Shotgun Metagenomics. *Comput. Struct. Biotechnol. J.* **2017**, *15*, 48–55. [[CrossRef](#)] [[PubMed](#)]
117. Mas-Lloret, J.; Obón-Santacana, M.; Ibáñez-Sanz, G.; Guinó, E.; Pato, M.L.; Rodríguez-Moranta, F.; Mata, A.; García-Rodríguez, A.; Pimenoff, V.N. Gut Microbiome Diversity Detected by High-Coverage 16S and Shotgun Sequencing of Matched Stool and Colon Biopsy Samples. *bioRxiv* **2019**, 742635. [[CrossRef](#)]
118. Bonin, N.; Doster, E.; Worley, H.; Pinnell, L.J.; Bravo, J.E.; Ferm, P.; Marini, S.; Prospero, M.; Noyes, N.; Morley, P.S.; et al. MEGARes and AMR++, v3.0: An Updated Comprehensive Database of Antimicrobial Resistance Determinants and an Improved Software Pipeline for Classification Using High-Throughput Sequencing. *Nucleic Acids Res.* **2022**, *51*, D744–D752. [[CrossRef](#)] [[PubMed](#)]
119. Mitchell, A.L.; Scheremetjov, M.; Denise, H.; Potter, S.; Tarkowska, A.; Qureshi, M.; Salazar, G.; Pesseat, S.; Boland, M.; Hunter, F.; et al. EBI Metagenomics in 2017: Enriching the Analysis of Microbial Communities, From Sequence Reads to Assemblies. *Nucleic Acids Res.* **2017**, *46*, D726–D735. [[CrossRef](#)] [[PubMed](#)]
120. Queyrel, M.; Prifti, E.; Templier, A.; Zucker, J.-D. Towards End-to-End Disease Prediction From Raw Metagenomic Data. *bioRxiv* **2020**. [[CrossRef](#)]
121. Oec, N.; Bono, H. Rapid Metagenomic Workflow Using Annotated 16S RNA Dataset. *BioHackrXiv* **2021**. [[CrossRef](#)]
122. Yang, C.; Chowdhury, D.; Zhang, Z.; Cheung, W.K.; Lu, A.; Bian, Z.; Zhang, L. A review of computational tools for generating metagenome-assembled genomes from metagenomic sequencing data. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 6301–6314. [[CrossRef](#)]
123. Xia, Y. Statistical normalization methods in microbiome data with application to microbiome cancer research. *Gut Microbes* **2023**, *15*, 2244139. [[CrossRef](#)]
124. Jennings, L.J.; Arcila, M.E.; Corless, C.; Kamel-Reid, S.; Lubin, I.M.; Pfeifer, J.; Temple-Smolkin, R.L.; Voelkerding, K.V.; Nikiforova, M.N. Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. *J. Mol. Diagn.* **2017**, *19*, 341–365. [[CrossRef](#)]
125. Shamim, K.; Sharma, J.; Dubey, S.K. Rapid and Efficient Method to Extract Metagenomic DNA From Estuarine Sediments. *3 Biotech.* **2017**. [[CrossRef](#)]
126. Ruocco, N.; Costantini, S.; Zupo, V.; Romano, G.; Ianora, A.; Fontana, A.; Costantini, M. High-Quality RNA Extraction From the Sea Urchin *Paracentrotus Lividus* Embryos. *PLoS ONE* **2017**. [[CrossRef](#)]
127. Crofts, T.S.; McFarland, A.; Hartmann, E.M. Mosaic Ends Tagmentation (METa) Assembly for Extremely Efficient Construction of Functional Metagenomic Libraries. *Msystems* **2021**. [[CrossRef](#)] [[PubMed](#)]

128. Peng, Z.; Zhu, X.; Wang, Z.; Yan, X.; Wang, G.; Meifang, T.; Jiang, A.; Kristiansen, K. Comparative Analysis of Sample Extraction and Library Construction for Shotgun Metagenomics. *Bioinform. Biol. Insights* **2020**. [[CrossRef](#)] [[PubMed](#)]
129. Palomares, M.-A.; Dalmasso, C.; Bonnet, E.; Derbois, C.; Brohard, S.; Ambroise, C.; Battail, C.; Deleuze, J.F.; Olasso, R. Comprehensive Analysis of RNA-seq Kits for Standard, Low and Ultra-Low Quantity Samples. *bioRxiv* **2019**. [[CrossRef](#)]
130. Hon, T.; Mars, K.; Young, G.; Tsai, Y.-C.; Karalius, J.W.; Landolin, J.M.; Maurer, N.; Kudrna, D.; Hardigan, M.A.; Steiner, C.C.; et al. Highly accurate long-read HiFi sequencing data for five complex genomes. *Sci. Data* **2020**, *7*, 399. [[CrossRef](#)] [[PubMed](#)]
131. Amarasinghe, S.L.; Su, S.; Dong, X.; Zappia, L.; Ritchie, M.E.; Gouil, Q. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol.* **2020**, *21*, 30. [[CrossRef](#)]
132. Orellana, L.H.; Krüger, K.; Sidhu, C.; Amann, R. Comparing Genomes Recovered From Time-Series Metagenomes Using Long-And Short-Read Sequencing Technologies. *Microbiome* **2023**. [[CrossRef](#)]
133. Browne, P.D.; Nielsen, T.K.; Kot, W.; Aggerholm, A.; Gilbert, M.T.P.; Puetz, L.; Rasmussen, M.; Zervas, A.; Hansen, L.H. GC bias affects genomic and metagenomic reconstructions, underrepresenting GC-poor organisms. *Gigascience* **2020**, *9*, gaaa008. [[CrossRef](#)]
134. Weissman, J.L.; Peras, M.; Barnum, T.P.; Fuhrman, J.A. Benchmarking Community-Wide Estimates of Growth Potential from Metagenomes Using Codon Usage Statistics. *mSystems* **2022**, *7*, e0074522. [[CrossRef](#)]
135. Sato, M.; Ogura, Y.; Nakamura, K.; Nishida, R.; Hayashi, M.; Hisatsune, J.; Sugai, M.; Itoh, T.; Hayashi, T. Comparison of the Sequencing Bias of Currently Available Library Preparation Kits for Illumina Sequencing of Bacterial Genomes and Metagenomes. *DNA Res.* **2019**, dsz017. [[CrossRef](#)]
136. Chouvarine, P.; Wiehlmann, L.; Moran Losada, P.; DeLuca, D.S.; Tümmler, B. Filtration and Normalization of Sequencing Read Data in Whole-Metagenome Shotgun Samples. *PLoS ONE* **2016**, *11*, e0165015. [[CrossRef](#)]
137. Brumfield, K.D.; Shanks, O.C.; Sivaganesan, M.; Hey, J.; Hasan, N.A.; Huq, A.; Colwell, R.R.; Leddy, M.B. Metagenomic Sequencing and Quantitative Real-Time PCR for Fecal Pollution Assessment in an Urban Watershed. *Front. Water* **2021**, *2021*, 626849. [[CrossRef](#)]
138. Mandal, S.D.; Panda, A.K.; Lalnunmawii, E.; Bisht, S.S.; Kumar, N.S. Illumina-Based Analysis of Bacterial Community in Khuangcherapuk Cave of Mizoram, Northeast India. *Genom. Data* **2015**. [[CrossRef](#)]
139. Fu, Y.; Wu, P.-H.; Beane, T.; Zamore, P.D.; Weng, Z. Elimination of PCR duplicates in RNA-seq and small RNA-seq using unique molecular identifiers. *BMC Genom.* **2018**, *19*, 531. [[CrossRef](#)]
140. Saheb Kashaf, S.; Almeida, A.; Segre, J.A.; Finn, R.D. Recovering prokaryotic genomes from host-associated, short-read shotgun metagenomic sequencing data. *Nat. Protoc.* **2021**, *16*, 2520–2541. [[CrossRef](#)]
141. Simner, P.J.; Miller, S.A.; Carroll, K.C. Understanding the Promises and Hurdles of Metagenomic Next-Generation Sequencing as a Diagnostic Tool for Infectious Diseases. *Clin. Infect. Dis.* **2017**, *66*, 778–788. [[CrossRef](#)]
142. Quick, J.; Grubaugh, N.D.; Pullan, S.T.; Claro, I.M.; Smith, A.D.; Gangavarapu, K.; Oliveira, G.; Robles-Sikisaka, R.; Rogers, T.F.; Beutler, N.; et al. Multiplex PCR Method for MinION and Illumina Sequencing of Zika and Other Virus Genomes Directly from Clinical Samples. *Nat. Protoc.* **2017**, *12*, 1261–1276. [[CrossRef](#)] [[PubMed](#)]
143. Charalampous, T.; Kay, G.L.; Richardson, H.; Aydin, A.; Baldan, R.; Jeanes, C.; Rae, D.; Grundy, S.; Turner, D.J.; Wain, J.; et al. Nanopore Metagenomics Enables Rapid Clinical Diagnosis of Bacterial Lower Respiratory Infection. *Nat. Biotechnol.* **2019**, *37*, 783–792. [[CrossRef](#)] [[PubMed](#)]
144. Buytaers, F.E.; Saltykova, A.; Denayer, S.; Verhaegen, B.; Vanneste, K.; Roosens, N.; Piérard, D.; Marchal, K.; Keersmaecker, S.C.J.D. A Practical Method to Implement Strain-Level Metagenomics-Based Foodborne Outbreak Investigation and Source Tracking in Routine. *Microorganisms* **2020**, *8*, 1191. [[CrossRef](#)] [[PubMed](#)]
145. Hendriksen, R.S.; Munk, P.; Njage, P.M.K.; Bunnik, B.A.D.v.; McNally, L.; Lukjancenka, O.; Röder, T.; Nieuwenhuijse, D.F.; Pedersen, S.K.; Kjeldgaard, J.S.; et al. Global Monitoring of Antimicrobial Resistance Based on Metagenomics Analyses of Urban Sewage. *Nat. Commun.* **2019**, *10*, 1124. [[CrossRef](#)] [[PubMed](#)]
146. Pehrsson, E.C.; Tsukayama, P.; Patel, S.; Mejía-Bautista, M.; Sosa-Soto, G.; Navarrete, K.M.; Calderón, M.; Cabrera, L.; Hoyos-Arango, W.; Bértoli, M.; et al. Interconnected Microbiomes and Resistomes in Low-Income Human Habitats. *Nature* **2016**, *533*, 212–216. [[CrossRef](#)]
147. Zuo, T.; Zhang, F.; Lui, G.; Yeoh, Y.K.; Li, A.Y.L.; Zhan, H.; Wan, Y.; Chung, A.C.K.; Cheung, C.P.; Chen, N.; et al. Alterations in Gut Microbiota of Patients With COVID-19 During Time of Hospitalization. *Gastroenterology* **2020**, *159*, 944–955.e8. [[CrossRef](#)]
148. Quince, C.; Walker, A.W.; Simpson, J.T.; Loman, N.J.; Segata, N. Shotgun metagenomics, from sampling to analysis. *Nat. Biotechnol.* **2017**, *35*, 833–844. [[CrossRef](#)]
149. McEvoy, C.R.E.; Semple, T.; Yellapu, B.; Choong, D.Y.H.; Xu, H.; Arnau, G.M.; Fellowes, A.; Fox, S.B. Improved Next-Generation Sequencing Pre-Capture Library Yields and Sequencing Parameters Using on-Bead PCR. *BioTechniques* **2020**, *68*, 48–51. [[CrossRef](#)]
150. Oechslin, C.P.; Lenz, N.; Liechti, N.; Ryter, S.; Agyeman, P.; Bruggmann, R.; Leib, S.L.; Beuret, C. Limited Correlation of Shotgun Metagenomics Following Host Depletion and Routine Diagnostics for Viruses and Bacteria in Low Concentrated Surrogate and Clinical Samples. *Front. Cell. Infect. Microbiol.* **2018**, *8*, 375. [[CrossRef](#)] [[PubMed](#)]
151. Espy, M.J.; Rys, P.N.; Wold, A.D.; Uhl, J.R.; Sloan, L.M.; Jenkins, G.D.; Ilstrup, D.M.; Cockerill, F.R.; Patel, R.; Rosenblatt, J.; et al. Detection of Herpes Simplex Virus DNA in Genital and Dermal Specimens by LightCycler PCR After Extraction Using the IsoQuick, MagNA Pure, and BioRobot 9604 Methods. *J. Clin. Microbiol.* **2001**, *39*, 2233–2236. [[CrossRef](#)] [[PubMed](#)]

152. Schorling, S.; Schalasta, G.; Enders, G.; Zauke, M. Quantification of Parvovirus B19 DNA Using COBAS AmpliPrep Automated Sample Preparation and LightCycler Real-Time PCR. *J. Mol. Diagn.* **2004**, *6*, 37–41. [[CrossRef](#)] [[PubMed](#)]
153. Pearlman, S.I.; Leelawong, M.; Richardson, K.A.; Adams, N.M.; Russ, P.K.; Pask, M.E.; Wolfe, A.E.; Wessely, C.; Haselton, F.R. Low-Resource Nucleic Acid Extraction Method Enabled by High-Gradient Magnetic Separation. *ACS Appl. Mater. Interfaces* **2020**, *12*, 12457–12467. [[CrossRef](#)] [[PubMed](#)]
154. Abdel-Latif, A.; Osman, G. Comparison of three genomic DNA extraction methods to obtain high DNA quality from maize. *Plant Methods* **2017**, *13*, 1. [[CrossRef](#)]
155. Li, X.; Bosch-Tijhof, C.J.; Wei, X.; de Soet, J.J.; Crielaard, W.; Loveren, C.V.; Deng, D.M. Efficiency of chemical versus mechanical disruption methods of DNA extraction for the identification of oral Gram-positive and Gram-negative bacteria. *J. Int. Med. Res.* **2020**, *48*, 300060520925594. [[CrossRef](#)] [[PubMed](#)]
156. Huq, T.; Khan, A.; Brown, D.; Dhayagude, N.; He, Z.; Ni, Y. Sources, production and commercial applications of fungal chitosan: A review. *J. Bioresour. Bioprod.* **2022**, *7*, 85–98. [[CrossRef](#)]
157. Ramírez, A.C.; Cailleau, G.; Fattou, M.; Dorador, C.; Junier, P. Diversity of Lysis-Resistant Bacteria and Archaea in the Polyextreme Environment of Salar De Huasco. *Front. Microbiol.* **2022**, *13*, 826117. [[CrossRef](#)]
158. Yasui, T.; Yanagida, T.; Shimada, T.; Otsuka, K.; Takeuchi, M.; Nagashima, K.; Rahong, S.; Naito, T.; Takeshita, D.; Yonese, A.; et al. Engineering Nanowire-Mediated Cell Lysis for Microbial Cell Identification. *ACS Nano* **2019**, *13*, 2262–2273. [[CrossRef](#)]
159. Ramírez, A.C.; Bregnard, D.; Junier, T.; Cailleau, G.; Dorador, C.; Bindschedler, S.; Junier, P. Assessment of Fungal Spores and Spore-Like Diversity in Environmental Samples by Targeted Lysis. *BMC Microbiol.* **2023**, *23*, 68. [[CrossRef](#)]
160. Nittala, P.V.K.; Hohreiter, A.; Linhard, E.R.; Dohn, R.; Mishra, S.; Konda, A.; Divan, R.; Guha, S.; Basu, A. Integration of Silicon Chip Microstructures for in-Line Microbial Cell Lysis in Soft Microfluidics. *Lab Chip* **2023**, *23*, 2327–2340. [[CrossRef](#)]
161. Zinter, M.; Mayday, M.; Ryckman, K.; Jelliffe-Pawlowski, L.; DeRisi, J. Towards precision quantification of contamination in metagenomic sequencing experiments. *Microbiome* **2019**, *7*, 62. [[CrossRef](#)]
162. Ogunbayo, A.E.; Sabiu, S.; Nyaga, M.M. Evaluation of extraction and enrichment methods for recovery of respiratory RNA viruses in a metagenomics approach. *J. Virol. Methods* **2023**, *314*, 114677. [[CrossRef](#)]
163. Zhang, L.; Fang, X.; Liao, H.; Zhang, Z.; Zhou, X.; Chen, Y.; Qiu, Q.; Li, S.C. A Comprehensive Investigation of Metagenome Assembly by Linked-Read Sequencing. *Microbiome* **2020**, *8*, 156. [[CrossRef](#)] [[PubMed](#)]
164. Zhou, F.; Xu, Y. cBar: A Computer Program to Distinguish Plasmid-Derived From Chromosome-Derived Sequence Fragments in Metagenomics Data. *Bioinformatics* **2010**, *26*, 2051–2052. [[CrossRef](#)] [[PubMed](#)]
165. Noguchi, H.; Park, J.H.; Takagi, T. MetaGene: Prokaryotic Gene Finding From Environmental Genome Shotgun Sequences. *Nucleic Acids Res.* **2006**, *34*, 5623–5630. [[CrossRef](#)] [[PubMed](#)]
166. Sharpton, T.J.; Riesenfeld, S.J.; Kembel, S.W.; Ladau, J.; O'Dwyer, J.P.; Green, J.L.; Eisen, J.A.; Pollard, K.S. PhylOTU: A High-Throughput Procedure Quantifies Microbial Community Diversity and Resolves Novel Taxa from Metagenomic Data. *PLoS Comput. Biol.* **2011**, *7*, e1001061. [[CrossRef](#)] [[PubMed](#)]
167. Thomas, J.C.; Oladeinde, A.; Kieran, T.J.; Finger, J.W.; Bayona-Vásquez, N.J.; Cartee, J.C.; Beasley, J.C.; Seaman, J.C.; McArthur, J.V.; Rhodes, O.E.; et al. Co-occurrence of Antibiotic, Biocide, and Heavy Metal Resistance Genes in Bacteria From Metal and Radionuclide Contaminated Soils at the Savannah River Site. *Microb. Biotechnol.* **2020**, *13*, 1179–1200. [[CrossRef](#)] [[PubMed](#)]
168. Gaulke, C.A.; Schmeltzer, E.R.; Dasenko, M.; Tyler, B.M.; Thurber, R.V.; Sharpton, T.J. Evaluation of the Effects of Library Preparation Procedure and Sample Characteristics on the Accuracy of Metagenomic Profiles. *mSystems* **2021**, *6*, e0044021. [[CrossRef](#)] [[PubMed](#)]
169. Heinicke, F.; Zhong, X.; Zucknick, M.; Breidenbach, J.; Sundaram, A.; Flâm, S.T.; Leithaug, M.; Dalland, M.; Farmer, A.; Henderson, J.M.; et al. Systematic Assessment of Commercially Available Low-Input miRNA Library Preparation Kits. *RNA Biol.* **2019**, *17*, 75–86. [[CrossRef](#)] [[PubMed](#)]
170. Wang, C.; Zhang, L.; Jiang, X.; Ma, W.; Geng, H.; Wang, X.; Li, M. Toward Efficient and High-Fidelity Metagenomic Data From Sub-Nanogram DNA: Evaluation of Library Preparation and Decontamination Methods. *BMC Biol.* **2022**, *20*, 225. [[CrossRef](#)] [[PubMed](#)]
171. Jones, M.B.; Highlander, S.K.; Anderson, E.L.; Li, W.; Dayrit, M.; Klitgord, N.; Fabani, M.M.; Seguritan, V.; Green, J.; Pride, D.T.; et al. Library Preparation Methodology Can Influence Genomic and Functional Predictions in Human Microbiome Research. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 14024–14029. [[CrossRef](#)] [[PubMed](#)]
172. Petersen, L.M.; Martin, I.W.; Moschetti, W.E.; Kershaw, C.; Tsongalis, G.J. Third-Generation Sequencing in the Clinical Laboratory: Exploring the Advantages and Challenges of Nanopore Sequencing. *J. Clin. Microbiol.* **2019**, *58*. [[CrossRef](#)]
173. Poulsen, C.S.; Ekstrøm, C.T.; Aarestrup, F.M.; Pamp, S.J. Library Preparation and Sequencing Platform Introduce Bias in Metagenomic-Based Characterizations of Microbiomes. *Microbiol. Spectr.* **2022**, *10*, e0009022. [[CrossRef](#)]
174. Linde, J.; Brangsch, H.; Hölzer, M.; Thomas, C.; Elschner, M.C.; Melzer, F.; Tomaso, H. Comparison of Illumina and Oxford Nanopore Technology for genome analysis of *Francisella tularensis*, *Bacillus anthracis*, and *Brucella suis*. *BMC Genom.* **2023**, *24*, 258. [[CrossRef](#)]
175. Chen, J.; Xu, F. Application of Nanopore Sequencing in the Diagnosis and Treatment of Pulmonary Infections. *Mol. Diagn. Ther.* **2023**, *27*, 685–701. [[CrossRef](#)]
176. Ghurye, J.S.; Cepeda-Espinoza, V.; Pop, M. Metagenomic Assembly: Overview, Challenges and Applications. *Yale J. Biol. Med.* **2016**, *89*, 353–362.

177. Gihawi, A.; Cardenas, R.; Hurst, R.; Brewer, D.S. Quality Control in Metagenomics Data. In *Metagenomic Data Analysis*; Springer: New York, NY, USA, 2023; pp. 21–54.
178. Roev, G.V.; Borisova, N.I.; Chistyakova, N.V.; Agletdinov, M.R.; Akimkin, V.G.; Khafizov, K. Unlocking the Viral Universe: Metagenomic Analysis of Bat Samples Using Next-Generation Sequencing. *Microorganisms* **2023**, *11*, 2532. [CrossRef]
179. Wang, L.; Ding, R.; He, S.; Wang, Q.; Zhou, Y. A pipeline for constructing reference genomes for large cohort-specific metagenome compression. *Microorganisms* **2023**, *11*, 2560. [CrossRef]
180. Mushtaq, S.; Khan, M.I.U.; Khan, M.T.; Lodhi, M.S.; Wei, D.Q. Novel mutations in structural proteins of dengue virus genomes. *J. Infect. Public Health* **2023**, *16*, 1971–1981. [CrossRef]
181. Child, H.T.; Airey, G.; Maloney, D.M.; Parker, A.; Wild, J.; McGinley, S.; Evens, N.; Porter, J.; Templeton, K.; Paterson, S. Comparison of metagenomic and targeted methods for sequencing human pathogenic viruses from wastewater. *mBio* **2023**, *14*, e0146823. [CrossRef] [PubMed]
182. Bjerg, J.J.; Lustermans, J.J.M.; Marshall, I.P.G.; Mueller, A.; Brokjær, S.; Thorup, C.B.; Tataru, P.; Schmid, M.; Wagner, M.; Nielsen, L.P.; et al. Cable Bacteria With Electric Connection to Oxygen Attract Flocks of Diverse Bacteria. *Nat. Commun.* **2023**, *14*, 1614. [CrossRef]
183. Kayani, M.U.R.; Huang, W.; Feng, R.; Chen, L. Genome-resolved metagenomics using environmental and clinical samples. *Brief. Bioinform.* **2021**, *22*, bbab030. [CrossRef] [PubMed]
184. Li, D.; Liu, C.-M.; Luo, R.; Sadakane, K.; Lam, T.W. MEGAHIT: An Ultra-Fast Single-Node Solution for Large and Complex Metagenomics Assembly via Succinct *de Bruijn* Graph. *Bioinformatics* **2015**, *31*, 1674–1676. [CrossRef] [PubMed]
185. Martinez-Hernandez, J.E.; Berrios, P.; Santibáñez, R.; Astroz, Y.C.; Sanchez, C.; Martin, A.J.; Trombert, A.N. First metagenomic analysis of the Andean condor (*Vultur gryphus*) gut microbiome reveals microbial diversity and wide resistome. *PeerJ* **2023**, *11*, e15235. [CrossRef] [PubMed]
186. Ma, Y.; Wu, N.; Zhang, T.; Li, Y.; Cao, L.; Zhang, P.; Zhang, Z.; Zhu, T.; Zhang, C. The microbiome, resistome, and their co-evolution in sewage at a hospital for infectious diseases in Shanghai, China. *Microbiol. Spectr.* **2023**, *12*, e0390023. [CrossRef] [PubMed]
187. Mendes, C.I.; Vila-Cerqueira, P.; Motro, Y.; Moran-Gilad, J.; Carriço, J.A.; Ramirez, M. LMAS: Evaluating metagenomic short de novo assembly methods through defined communities. *GigaScience* **2023**, *12*, giac122. [CrossRef] [PubMed]
188. Cheung, M.K.; Ng, R.W.Y.; Lai, C.K.C.; Zhu, C.; Au, E.T.K.; Yau, J.W.K.; Li, C.; Wong, H.C.; Wong, B.C.K.; Kwok, K.O.; et al. Alterations in faecal microbiome and resistome in Chinese international travellers: A metagenomic analysis. *J. Travel Med.* **2023**, *30*, taad027. [CrossRef] [PubMed]
189. Yorke, S.; Shea, T.; Cuomo, C.A.; Walker, B.J.; LaRocque, R.C.; Manson, A.L.; Earl, A.M.; Worby, C.J. Comparison of long- and short-read metagenomic assembly for low-abundance species and resistance genes. *Brief. Bioinform.* **2023**, *24*, bbad050. [CrossRef] [PubMed]
190. Diao, Z.; Zhang, Y.; Chen, Y.; Han, Y.; Chang, L.; Ma, Y.; Feng, L.; Huang, T.; Zhang, R.; Li, J. Assessing the Quality of Metagenomic Next-Generation Sequencing for Pathogen Detection in Lower Respiratory Infections. *Clin. Chem.* **2023**, *69*, 1038–1049. [CrossRef]
191. Adekoya, A.E.; Kargbo, H.A.; Ibberson, C.B. Defining microbial community functions in chronic human infection with metatranscriptomics. *mSystems* **2023**, *8*, e00593-23. [CrossRef] [PubMed]
192. Blanco-Míguez, A.; Beghini, F.; Cumbo, F.; McIver, L.J.; Thompson, K.N.; Zolfo, M.; Manghi, P.; Dubois, L.; Huang, K.D.; Thomas, A.M. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nat. Biotechnol.* **2023**, *41*, 1633–1644. [CrossRef] [PubMed]
193. Martin-Cuadrado, A.-B.; Rubio-Portillo, E.; Antón, J. Community Shifts in the Coral *Oculina Patagonica* Holobiont in Response to Confinement, Temperature and *Vibrio* Infections. 2024. Available online: <https://www.researchsquare.com/article/rs-3893459/v1> (accessed on 1 February 2024).
194. Vigil, K.; Aw, T.G. Comparison of de novo assembly using long-read shotgun metagenomic sequencing of viruses in fecal and serum samples from marine mammals. *Front. Microbiol.* **2023**, *14*, 1248323. [CrossRef]
195. van Boheemen, S.; van Rijn, A.L.; Pappas, N.; Carbo, E.C.; Vorderman, R.H.P.; Sidorov, I.; van't Hof, P.J.; Mei, H.; Claas, E.C.J.; Kroes, A.C.M.; et al. Retrospective Validation of a Metagenomic Sequencing Protocol for Combined Detection of RNA and DNA Viruses Using Respiratory Samples from Pediatric Patients. *J. Mol. Diagn.* **2020**, *22*, 196–207. [CrossRef]
196. de Vries, J.J.C.; Brown, J.R.; Couto, N.; Beer, M.; Le Mercier, P.; Sidorov, I.; Papa, A.; Fischer, N.; Oude Munnink, B.B.; Rodriguez, C.; et al. Recommendations for the introduction of metagenomic next-generation sequencing in clinical virology, part II: Bioinformatic analysis and reporting. *J. Clin. Virol.* **2021**, *138*, 104812. [CrossRef]
197. Dulanto Chiang, A.; Dekker, J.P. From the pipeline to the bedside: Advances and challenges in clinical metagenomics. *J. Infect. Dis.* **2020**, *221*, S331–S340. [CrossRef]
198. Vilsker, M.; Moosa, Y.; Nooij, S.; Fonseca, V.; Ghysens, Y.; Dumon, K.; Pauwels, R.; Alcantara, L.C.; Vanden Eynden, E.; Vandamme, A.-M.; et al. Genome Detective: An automated system for virus identification from high-throughput sequencing data. *Bioinformatics* **2018**, *35*, 871–873. [CrossRef]
199. Shao, L.; Liao, J.; Qian, J.; Chen, W.; Fan, X. MetaGeneBank: A Standardized Database to Study Deep Sequenced Metagenomic Data From Human Fecal Specimen. *BMC Microbiol.* **2021**, *21*, 263. [CrossRef]
200. Páez-Espino, D.; Eloie-Fadrosch, E.A.; Pavlopoulos, G.A.; Thomas, A.D.; Huntemann, M.; Mikhailova, N.; Rubin, E.M.; Ivanova, N.; Kyrpides, N.C. Uncovering Earth's Virome. *Nature* **2016**, *536*, 425–430. [CrossRef] [PubMed]

201. Sichtig, H.; Minogue, T.D.; Yan, Y.; Stefan, C.P.; Hall, A.T.; Tallon, L.J.; Sadzewicz, L.; Nadendla, S.; Klimke, W.; Hatcher, E.L.; et al. FDA-ARGOS Is a Database With Public Quality-Controlled Reference Genomes for Diagnostic Use and Regulatory Science. *Nat. Commun.* **2019**, *33*, 3313. [CrossRef] [PubMed]
202. Parks, D.H.; Rigato, F.; Vera-Wolf, P.; Krause, L.; Hugenholtz, P.; Tyson, G.W.; Wood, D. Evaluation of the Microba Community Profiler for Taxonomic Profiling of Metagenomic Datasets From the Human Gut Microbiome. *Front. Microbiol.* **2021**, *12*, 643682. [CrossRef] [PubMed]
203. OSF. Metagenomics Benchmarking. Available online: <https://doi.org/10.17605/OSF.IO/RE4PD> (accessed on 20 January 2024).
204. Chen, Y.; Fan, L.; Chai, Y.-h.; Xu, J. Advantages and Challenges of Metagenomic Sequencing for the Diagnosis of Pulmonary Infectious Diseases. *Clin. Respir. J.* **2022**, *16*, 646–656. [CrossRef] [PubMed]
205. Ruppé, É.; Cherkaoui, A.; Lazarević, V.; Emonet, S.; Schrenzel, J. Establishing Genotype-to-Phenotype Relationships in Bacteria Causing Hospital-Acquired Pneumonia: A Prelude to the Application of Clinical Metagenomics. *Antibiotics* **2017**, *6*, 30. [CrossRef] [PubMed]
206. Goldberg, B.; Sichtig, H.; Geyer, C.N.; Ledebor, N.A.; Weinstock, G.M. Making the Leap From Research Laboratory to Clinic: Challenges and Opportunities for Next-Generation Sequencing in Infectious Disease Diagnostics. *mBio* **2015**, *6*, 10–128. [CrossRef] [PubMed]
207. Edgeworth, J.D. Respiratory Metagenomics: Route to Routine Service. *Curr. Opin. Infect. Dis.* **2023**, *36*, 115–123. [CrossRef]
208. Govender, K.; Street, T.; Sanderson, N.; Eyre, D.W. Metagenomic Sequencing as a Pathogen-Agnostic Clinical Diagnostic Tool for Infectious Diseases: A Systematic Review and Meta-Analysis of Diagnostic Test Accuracy Studies. *J. Clin. Microbiol.* **2021**, *59*, 10–128. [CrossRef]
209. Blauwkamp, T.A.; Thair, S.; Rosen, M.J.; Blair, L.; Lindner, M.S.; Vilfan, I.D.; Kawli, T.; Christians, F.C.; Venkatasubrahmanyam, S.; Wall, G.D.; et al. Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. *Nat. Microbiol.* **2019**, *4*, 663–674. [CrossRef]
210. Rehm, H.L.; Bale, S.J.; Bayrak-Toydemir, P.; Berg, J.S.; Brown, K.K.; Deignan, J.L.; Friez, M.J.; Funke, B.H.; Hegde, M.R.; Lyon, E. ACMG clinical laboratory standards for next-generation sequencing. *Genet. Med. Off. J. Am. Coll. Med. Genet.* **2013**, *15*, 733–747. [CrossRef]
211. Diao, Z.; Lai, H.; Han, D.; Yang, B.; Zhang, R.; Li, J. Validation of a Metagenomic Next-Generation Sequencing Assay for Lower Respiratory Pathogen Detection. *Microbiol. Spectr.* **2023**, *11*, e0381222. [CrossRef]
212. FDA Guidance Documents. Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests—Guidance for Industry and FDA Staff. Available online: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/statistical-guidance-reporting-results-studies-evaluating-diagnostic-tests-guidance-industry-and-fda> (accessed on 4 January 2024).
213. Mee, E.T.; Preston, M.D.; Minor, P.D.; Schepelmann, S.; Huang, X.; Nguyen, J.; Wall, D.; Hargrove, S.; Fu, T.; Xu, G.; et al. Development of a candidate reference material for adventitious virus detection in vaccine and biologicals manufacturing by deep sequencing. *Vaccine* **2016**, *34*, 2035–2043. [CrossRef]
214. Sangwan, N.; Xia, F.; Gilbert, J.A. Recovering Complete and Draft Population Genomes From Metagenome Datasets. *Microbiome* **2016**, *4*, 8. [CrossRef] [PubMed]
215. Lai, B.; Wang, F.; Wang, X.; Duan, L.; Zhu, H. InteMAP: Integrated Metagenomic Assembly Pipeline for NGS Short Reads. *BMC Bioinform.* **2015**, *16*, 66. [CrossRef] [PubMed]
216. Hua, K.; Zhang, X. Estimating the Total Genome Length of a Metagenomic Sample Using K-Mers. *BMC Genom.* **2019**, *20*, 183. [CrossRef] [PubMed]
217. Norling, M.; Lindsjö, O.K.; Gourel, H.; Bongcam-Rudloff, E.; Hayer, J. MetLab: An in Silico Experimental Design, Simulation and Analysis Tool for Viral Metagenomics Studies. *PLoS ONE* **2016**, *11*, e0160334. [CrossRef] [PubMed]
218. Zaheer, R.; Noyes, N.; Polo, R.O.; Cook, S.R.; Marinier, E.; Domselaar, G.V.; Belk, K.E.; Morley, P.S.; McAllister, T.A. Impact of Sequencing Depth on the Characterization of the Microbiome and Resistome. *Sci. Rep.* **2018**, *8*, 5890. [CrossRef] [PubMed]
219. Chen, L.X.; Anantharaman, K.; Shaiber, A.; Eren, A.M.; Banfield, J.F. Accurate and complete genomes from metagenomes. *Genome Res.* **2020**, *30*, 315–333. [CrossRef] [PubMed]
220. Am, R.; Ar, R.; Rg, M.; Seah, C.; Nr, Y.; Dr, M.; Ag, M.; Ph, S.; Coburn, B. Performance Characteristics of Next-Generation Sequencing for Antimicrobial Resistance Gene Detection in Genomes and Metagenomes. *bioRxiv* **2021**. [CrossRef]
221. Pereira-Marques, J.; Hout, A.v.d.; Ferreira, R.M.; Weber, M.; Pinto-Ribeiro, I.; Doorn, L.J.v.; Knetsch, C.W.; Figueiredo, C. Impact of Host DNA and Sequencing Depth on the Taxonomic Resolution of Whole Metagenome Sequencing for Microbiome Analysis. *Front. Microbiol.* **2019**, *10*, 1277. [CrossRef]
222. Mattocks, C.J.; Morris, M.A.; Matthijs, G.; Swinnen, E.; Corveleyn, A.; Dequeker, E.; Müller, C.R.; Pratt, V.; Wallace, A. A standardized framework for the validation and verification of clinical molecular genetic tests. *Eur. J. Hum. Genet.* **2010**, *18*, 1276–1288. [CrossRef]
223. Ounit, R.; Wanamaker, S.; Close, T.J.; Lonardi, S. CLARK: Fast and Accurate Classification of Metagenomic and Genomic Sequences Using Discriminative K-Mers. *BMC Genom.* **2015**, *16*, 236. [CrossRef]
224. Wylie, T.N.; Wylie, K.M.; Herter, B.N.; Storch, G.A. Enhanced virome sequencing using targeted sequence capture. *Genome Res.* **2015**, *25*, 1910–1920. [CrossRef]

225. Hasan, M.R.; Rawat, A.; Tang, P.; Jithesh, P.V.; Thomas, E.; Tan, R.; Tilley, P. Depletion of human DNA in spiked clinical specimens for improvement of sensitivity of pathogen detection by next-generation sequencing. *J. Clin. Microbiol.* **2016**, *54*, 919–927. [[CrossRef](#)]
226. Breitwieser, F.P.; Baker, D.; Salzberg, S.L. KrakenUniq: Confident and Fast Metagenomics Classification Using Unique K-Mer Counts. *Genome Biol.* **2018**, *19*, 198. [[CrossRef](#)]
227. Greninger, A.L.; Chen, E.C.; Sittler, T.; Scheinerman, A.; Roubinian, N.; Yu, G.; Kim, E.; Pillai, D.R.; Guyard, C.; Mazzulli, T.; et al. A Metagenomic Analysis of Pandemic Influenza a (2009 H1N1) Infection in Patients From North America. *PLoS ONE* **2010**, *5*, e13381. [[CrossRef](#)]
228. Orellana, L.H.; Rodriguez-R, L.M.; Konstantinidis, K.T. ROcker: Accurate Detection and Quantification of Target Genes in Short-Read Metagenomic Data Sets by Modeling Sliding-Window Bitscores. *Nucleic Acids Res.* **2016**, *45*, e14. [[CrossRef](#)]
229. Biney-Assan, T.; Kron, M. Molecular Microbiology in Clinical Practice: Current and Future Applications. *Al-Kindy Coll. Med. J.* **2022**, *18*, 167–172. [[CrossRef](#)]
230. Altermann, E.; Tegetmeyer, H.E.; Chanyi, R.M. The Evolution of Bacterial Genome Assemblies—Where Do We Need to Go Next? *Microbiome Res. Rep.* **2022**, *1*, 15. [[CrossRef](#)]
231. Luan, Y.; Hu, H.; Liu, C.; Chen, B.; Liu, X.; Xu, Y.; Luo, X.; Chen, J.; Ye, B.; Huang, F.; et al. A proof-of-concept study of an automated solution for clinical metagenomic next-generation sequencing. *J. Appl. Microbiol.* **2021**, *131*, 1007–1016. [[CrossRef](#)]
232. PacBio. *PacBio Announces Collaboration with Leading Library Preparation Automation Partners*; PacBio: Menlo Park, CA, USA, 2023.
233. Diao, Z.; Han, D.; Zhang, R.; Li, J. Metagenomics next-generation sequencing tests take the stage in the diagnosis of lower respiratory tract infections. *J. Adv. Res.* **2022**, *38*, 201–212. [[CrossRef](#)]
234. Yen, S.; Johnson, J. Metagenomics: A Path to Understanding the Gut Microbiome. *Mamm. Genome* **2021**, *32*, 282–296. [[CrossRef](#)]
235. Jagadeesan, B.; Gerner-Smidt, P.; Allard, M.W.; Leuillet, S.; Winkler, A.; Xiao, Y.; Chaffron, S.; Van Der Vossen, J.; Tang, S.; Katase, M.; et al. The Use of Next Generation Sequencing for Improving Food Safety: Translation Into Practice. *Food Microbiol.* **2018**, *79*, 96–115. [[CrossRef](#)]
236. Angers-Loustau, A.; Petrillo, M.; Bengtsson-Palme, J.; Berendonk, T.U.; Blais, B.W.; Chan, K.G.; Coque, T.M.; Hammer, P.; Heß, S.; Kagkli, D.M.; et al. The Challenges of Designing a Benchmark Strategy for Bioinformatics Pipelines in the Identification of Antimicrobial Resistance Determinants Using Next Generation Sequencing Technologies. *F1000Research* **2018**, *7*, 459. [[CrossRef](#)]
237. Satam, H.; Joshi, K.; Mangrolia, U.; Waghoo, S.; Zaidi, G.; Rawool, S.; Thakare, R.P.; Banday, S.; Mishra, A.K.; Das, G.; et al. Next-Generation Sequencing Technology: Current Trends and Advancements. *Biology* **2023**, *12*, 997. [[CrossRef](#)]
238. Shi, W.; Qi, H.; Sun, Q.; Fan, G.; Liu, S.; Wang, J.; Zhu, B.; Liu, H.; Zhao, F.; Wang, X.; et al. gcMeta: A Global Catalogue of Metagenomics platform to support the archiving, standardization and analysis of microbiome data. *Nucleic Acids Res.* **2019**, *47*, D637–D648. [[CrossRef](#)]
239. Chowdhury, A.S.; Call, D.R.; Broschat, S.L. PARGT: A Software Tool for Predicting Antimicrobial Resistance in Bacteria. *Sci. Rep.* **2020**, *10*, 11033. [[CrossRef](#)]
240. Erdem, G.; Kaptan, I.; Sharma, H.; Kumar, A.; Aylward, S.C.; Kapoor, A.; Shimamura, M. Cerebrospinal Fluid Analysis for Viruses by Metagenomic Next-Generation Sequencing in Pediatric Encephalitis: Not Yet Ready for Prime Time? *J. Child Neurol.* **2020**, *36*, 350–356. [[CrossRef](#)]
241. Field, D.; Amaral-Zettler, L.; Cochrane, G.; Cole, J.R.; Dawyndt, P.; Gilbert, J.A.; Glöckner, F.O.; Hirschman, L.; Karsch-Mizrachi, I.; Klenk, H.P.; et al. The Genomic Standards Consortium. *PLoS Biol.* **2011**, *9*, e1001088. [[CrossRef](#)]
242. Kasmanas, J.C.; Bartholomäus, A.; Corrêa, F.B.; Tal, T.; Jehmlich, N.; Herberth, G.; von Bergen, M.; Stadler, P.F.; de Leon Ferreira de Carvalho, A.C.P.; da Rocha, U.N. HumanMetagenomeDB: A Public Repository of Curated and Standardized Metadata for Human Metagenomes. *Nucleic Acids Res.* **2020**, *49*, D743–D750. [[CrossRef](#)]
243. Posey, J.E.; Harel, T.; Liu, P.; Rosenfeld, J.A.; James, R.A.; Akdemir, Z.H.C.; Walkiewicz, M.; Bi, W.; Xiao, R.; Ding, Y.; et al. Resolution of Disease Phenotypes Resulting From Multilocus Genomic Variation. *N. Engl. J. Med.* **2017**, *376*, 21–31. [[CrossRef](#)]
244. Zhu, X.; Yan, S.; Yuan, F.; Wan, S. The Applications of Nanopore Sequencing Technology in Pathogenic Microorganism Detection. *Can. J. Infect. Dis. Med. Microbiol.* **2020**, *2020*, 6675206. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.